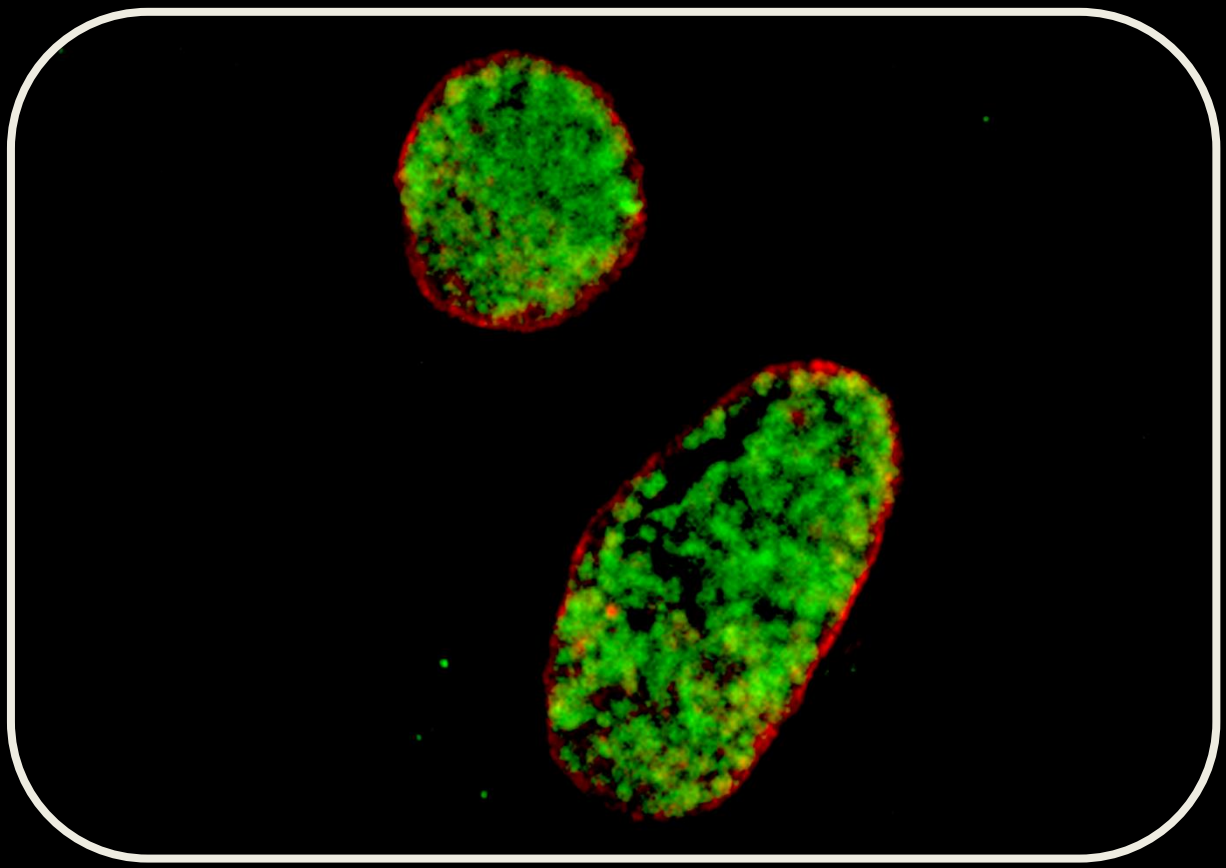
Functional analyses of inclusion membrane proteins of *Chlamydia trachomatis*

Filipe Manuel Baeta da Silva Almeida



Dissertation presented to obtain the Ph.D degree in Biology

Instituto de Tecnologia Química e Biológica António Xavier | Universidade Nova de Lisboa

Oeiras,
June, 2016

UNIVERSIDADE NOVA DE LISBOA

Oeiras, June, 2016

Functional analyses of inclusion membrane proteins of *Chlamydia trachomatis*

Filipe Manuel Baeta
da Silva Almeida

ITQB-UNL | Av. da República, 2780-157 Oeiras, Portugal
Tel (+351) 214 469 100 | Fax (+351) 214 411 277

www.itqb.unl.pt

Functional analyses of inclusion membrane proteins of *Chlamydia trachomatis*

Filipe Manuel Baeta da Silva Almeida

Dissertation presented to obtain the Ph.D degree in Biology

Instituto de Tecnologia Química e Biológica António Xavier | Universidade Nova de Lisboa

Oeiras, June, 2016

First Edition: March 2016
Second Edition: June 2016

Financial support from Fundação para a Ciência e a Tecnologia through the fellowship SFRH/BD/73545/2010 awarded to Filipe Almeida

(From left to right: Dr. Miguel Teixeira, Dr. João Paulo Gomes, Dr. Luís Jaime Mota, Filipe Almeida, Dr. Sandra Sousa, Dr. Richard Hayward, and Dr. Isabel Sá Nogueira)

- Supervisor:

Dr. Luís Jaime Mota, Head of the Infection Biology Lab at UCIBIO, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Monte da Caparica, Portugal

- President of the Juri:

Dr. Miguel Teixeira, Instituto de Tecnologia Química e Biológica, Oeiras, Portugal

- Juri:

Richard Hayward, Head of the Molecular Cell Biology and Bacterial Pathogenesis Lab at the Institute of and Structural of Molecular Biology, University of London, United Kingdom

Dr. Sandra Sousa, Assistant Researcher of the Molecular Microbiology Lab at the Instituto de Investigação e Inovação em Saude, Oporto, Portugal

Dr. Isabel Sá Nogueira, Head of the Microbial Genetics Lab at UCIBO, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Monte da Caparica, Portugal

Dr. João Paulo Gomes, Head of the Unidade de Investigação e Desenvolvimento at the Instituto Nacional de Saúde Dr. Ricardo Jorge, Lisbon, Portugal

Acknowledgments

First of all, I want to deeply thank my supervisor Jaime Mota, for his scientific rigor, expertise, patience, and motivation. Being constantly challenged by him to discuss the work and to search for new ideas helped to improve not only the thesis, but my experience as a scientific researcher. He has been very supportive for almost 8 years, since I joined him at the very beginning of the Infection Biology Lab at ITQB.

Besides my supervisor, I would like to acknowledge the members of my PhD committee Isabel Sá-Nogueira and João Paulo Gomes, for their insightful comments and encouragement throughout the thesis.

I want to acknowledge the institutions that made this thesis possible, by providing me all the necessary conditions and education: the ITQB, where I started this work; INSA, where I did some collaborations; FCT-NOVA, where I did the final experiments of the thesis; IGC and CEDOC for the microscopy facilities; and Fundação para a Ciência e Tecnologia for funding the project and my fellowship.

I want to give a special thanks to the past and the current staff of the IB lab Beatriz Costa, Catarina Milho, Inês Pereira, Irina Franco, Joana Bugalhão, Lia Domingues, Maria Cunha, Nuno Charro, and Sara Pais, with whom I shared the lab over these last years. Their great companionship and extensive discussions brought great value to the thesis, and to me as a person.

I would like to thank the people of the department of infectious diseases at INSA, namely, Maria José Borrego, João Paulo Gomes, Vítor Borges, Rita Ferreira, Carlos Florindo, and Vera Damião, with whom I spent a

couple of months during some collaborations with my lab. Part of the thesis was done there and I learnt a lot from their experience.

I want to thank the colleagues from the DCV, who welcomed us and let us be a part of the DCV community when we moved from ITQB to FCT-NOVA. Particularly, to Raquel Portela, Cynthia Barroco, Bárbara Gonçalves, Lia Godinho, and Mário Ferreira.

I thank other colleagues and staff from ITQB: Gonçalo Poças, who was my running partner a couple of times; the people of InTeraQB, with whom I shared a lot of fun moments while organizing scientific seminars and parties; Inês Grilo, Joana Rolo, Carina Valente, and Débora Tavares, with whom I spent the three months of PhD education; and the ITQB Science and Communication Office for all the support in the organization of the "Dia Aberto ITQB" initiatives.

Agradeço aos meus colegas do IST Ana Raposo, Joana Correia, Marta Franco e Vítor Farinha, com quem ainda mantenho contato. É sempre bom relembrar os tempos de faculdade.

Aos meus amigos de longa data, que me incentivaram a acabar a tese de doutoramento. Destaco o Tiago Trigo pelos belos concertos, o Rafael Possolo pelos jogos da bola, o Nuno pela emoção dada aos jogos de Agricola, e à Liliana pela amizade.

Por fim, mas no topo da minha lista, agradeço à minha família por todo o apoio ao longo da minha vida. À minha mãe e ao meu pai por terem feito os possíveis e os impossíveis para eu poder estar onde estou hoje e ao meu irmão por me ter sempre acompanhado nesta viagem. À minha avó Arminda, que apesar do seu estado de saúde atual, sempre me contagiou com a sua boa disposição, e aos meus queridos sobrinhos, Manuel e Duarte, pelas horas infindáveis de entretenimento e brincadeira que me proporcionam.

Thesis Publications

Almeida, F., Borges, V., Ferreira, R., Borrego, M.J., Gomes, J.P., and Mota, L.J. (2012) "Polymorphisms in Inc proteins and differential expression of *inc* genes among *Chlamydia trachomatis* strains correlate with invasiveness and tropism of lymphogranuloma venereum isolates". J Bacteriol **194**: 6574–6585.

da Cunha, M., Milho, C., Almeida, F., Pais, S.V., Borges, V., Maurício, R., Borrego, M. J., Gomes, J.P., Mota, L.J. (2014) "Identification of type III secretion substrates of *Chlamydia trachomatis* using *Yersinia enterocolitica* as a heterologous system". BMC Microbiol **14**: 40.

Almeida, F, Pais, S.V., Mota, L.J. "*Chlamydia trachomatis* inclusion membrane protein CT288 interacts with the host cell centrosomal protein CCDC146 (manuscript in preparation).

Abstract

Chlamydia trachomatis belongs to a group of highly related obligate intracellular bacteria. It is the etiological agent of important human sexually transmitted infections and trachoma. *C. trachomatis* strains can be divided in the trachoma biovar, which includes strains that cause non-invasive infections of conjunctival and genital epithelial cells (ocular and urogenital strains, respectively), and the lymphogranuloma venereum (LGV) biovar, which is responsible for a more invasive disease, and includes strains that are able to infect macrophages and disseminate into lymph nodes (LGV strains). The determinants of the different types of infection (invasive or non-invasive) and tissue tropism (eyes, genitals, and lymph nodes) are only partially studied. Furthermore, it is known that the genomes of *C. trachomatis* strains are nearly identical at the DNA level (>98%). Therefore, these differences must be explained by small nucleotide variations on a relatively low number of genes, which could lead either to proteins with disease group-specific amino acids or to differential gene expression.

Throughout the developmental cycle, *C. trachomatis* resides and replicates within the host cell inside a vacuole (an inclusion). It uses a type III secretion system (T3SS) to manipulate the host cell and benefit from the host supplies. The inclusion membrane (Inc) proteins are an important family of type III secretion (T3S) effectors. These proteins have a characteristic bi-lobed hydrophobic domain, likely responsible for targeting them to the membrane of the inclusion. Due to their particular localization, Inc proteins are obvious candidates to participate in the cross-talk between the bacteria and the host cell. Several Incs were already implicated in the manipulation of the host cytoskeleton dynamics, and vesicular and non-vesicular transport, however, the

function of the majority of them is still unknown. Aditionally, not all Incs were yet described as being T3S substrates, nor the localization in the inclusion membrane was described for all of them.

In this thesis we aimed to gain insights into the molecular function of Inc proteins. First we tested if the presence of a T3S signal could be used as additional evidence (other than the bi-lobed hydrophobic motif) that an Inc protein will localize at the inclusion membrane. We used *Yersinia enterocolitica* as heterologous host to identify T3S signals in a set of 48 *C. trachomatis* known (possessing the hydrophobic motif and described to be localized at the inclusion membrane) and putative (possessing the hydrophobic motif but not yet described to be localized at the inclusion membrane) Inc proteins. We identified a T3S signal in ~80% of the Incs analyzed, indicating that they are T3S substrates. However, we did not detect a T3S signal in a few known Inc proteins, and therefore, to validate a protein as an Inc protein, additional immunofluorescence studies in *Chlamydia*-infected cells are needed to confirm its localization at the inclusion membrane.

Given the particular localization of Inc proteins at the inclusion membrane, we next assessed if Inc proteins could participate in the tissue tropism and type of infection associated with *C. trachomatis*. For that, we performed phylogenetic, molecular evolution, and gene expression analyses, focusing on the same set of 48 Inc proteins retrieved from the sequenced genomes of 51 *C. trachomatis* strains. We discovered small differences in the amino acid sequence of a subgroup of Inc proteins and in the expression of *inc* genes that could contribute to the specific tissue tropism and type of infection associated with LGV strains.

To test the hypothesis that Incs contribute to tissue tropism and type of infection, we aimed to gain insight into the molecular function of the Inc

proteins highlighted in the previous study. For this, we used yeast-two hybrid screening with Inc proteins CT228, CT249, and CT288 as baits for a mammalian cDNA library. Among the putative candidates found, we characterized in more detail the interaction between CT288 and the coiled coil domain containing protein 146 (CCDC146). CCDC146 is a human protein of unknown function that localizes at the centrosome. We used different biochemical methods to validate the interaction between CT288 and CCDC146 and immunolocalization studies in *Chlamydia*-infected cells to study their localization. We show that during infection by *C. trachomatis*, CCDC146 co-localizes with CT288 at the inclusion membrane. It is known that the centrosome is targeted by *C. trachomatis* throughout the infectious cycle. Moreover, a few host centrosomal proteins were already identified as host cell targets of Inc proteins. We hypothesize that the interaction between CT288 and the host cell centrosomal protein CCDC146 might also be relevant for the cross-talk between the inclusion and the centrosome.

In conclusion, we gained novel insights into the pathogenesis of infections by *C. trachomatis*, particularly regarding proteins involved in tissue tropism and different types of infection. We also discovered a new potential host cell target for an Inc protein.

Resumo

Chlamydia trachomatis pertence a um grupo de bactérias intracelulares obrigatórias. É o agente etiológico de doenças de grande importância clínica como o tracoma e doenças sexualmente transmissíveis. As estirpes de *C. trachomatis* podem ser divididas em dois biovares: o biovar tracoma, que inclui as estirpes que causam infeções não invasivas do tecido epitelial da conjuntiva e do trato genital (estirpes oculares e genitais), e o biovar linfogranuloma venéreo (LGV), responsável por uma doença mais invasiva e que inclui estirpes que são capazes de infetar macrófagos e disseminar para os nódulos linfáticos. Os determinantes envolvidos nos diferentes tipos de infeção (não invasiva e invasiva) e nos diferentes tropismos celulares (olhos, genitais e nódulos linfáticos) estão apenas parcialmente estudados. Sabe-se ainda que os genomas das estirpes de *C. trachomatis* são quase idênticos entre si (>98% de semelhança ao nível do DNA). Assim, as diferenças no tipo de infeção e no tropismo celular devem ser explicadas por pequenas variações nucleotídicas num número restrito de genes, que podem dar origem a proteínas com aminoácidos específicos de cada grupo de doença e/ou a expressão diferenciada dos respetivos genes.

Durante o seu ciclo infecioso, *C. trachomatis* instala-se e replica dentro da célula hospedeira num vacúolo (designado por inclusão), e usa um sistema de secreção do tipo III (T3SS) para manipular a célula hospedeira. Um grupo importante de efetores do T3SS são as proteínas localizadas na membrana do inclusão (proteínas Inc). Estas proteínas são caraterizadas por possuir um domínio bilobal hidrofóbico, possivelmente capaz de as direcionar para a membrana do inclusão. Dada a sua particular localização, as proteínas Inc são fortes candidatas

a serem os fatores bacterianos responsáveis pela interação entre a bactéria e a célula hospedeira. Algumas proteínas Inc foram já descritas como estando envolvidas na manipulação de alguns processos do hospedeiro como a dinâmica do citosqueleto, e o transporte vesicular e não vesicular. No entanto, a função da maioria das proteínas Inc ainda é desconhecida. Além disso, nem todas as proteínas Inc foram já descritas como sendo efetores do T3SS, bem como a localização na membrana do inclusão ainda não foi confirmada para todas as Incs.

O objetivo desta tese foi o de aumentar o nosso conhecimento acerca da função molecular das proteínas Inc. Primeiro, testámos se a presença de um sinal de secreção do tipo III (T3S) poderia ser usada como evidência adicional (além do domínio bilobal caraterístico) de que uma proteína Inc se irá localizar na membrana do inclusão. Para isso, usámos *Yersinia enterocolitica* como sistema heterólogo para identificar sinais de T3S num conjunto de 48 proteínas Inc de *C. trachomatis*, algumas já conhecidas (sabe-se que têm o domínio bilobal e a sua localização na membrana do inclusão já foi descrita), outras apenas putativas (sabe-se apenas que têm o domínio bilobal). Identificámos um sinal de T3S em 80% das Incs, confirmando a hipótese de que são substratos do T3SS. No entanto, não conseguimos identificar um sinal de T3S em algumas das Incs já conhecidas. Isto significa que a presença de um sinal de T3S não é suficiente para validar uma proteína como sendo uma Inc. Experiências adicionais de microscopia de imunofluorescência em células infetadas por *Chlamydia* são necessárias para confirmar a localização das proteínas na membrana do inclusão.

Uma vez que as proteínas Inc se localizam na membrana do inclusão, a seguir avaliámos se estas proteínas podem contribuir para o tropismo celular e tipo de infeção associados às estirpes de *C. trachomatis*. Para isso, fizemos análises de filogenia, de evolução molecular, e de

expressão génica no mesmo grupo de 48 proteínas Inc obtidas dos genomas sequenciados de 51 estirpes de *C. trachomatis*. Descobrimos que pequenas diferenças ao nível da sequência de aminoácidos de um subgrupo de proteínas Inc, e ao nível da expressão de alguns genes que as codificam, podem contribuir para o tropismo e tipo de infeção específico associado às estirpes LGV.

Para conseguir testar a hipótese de que as Incs contribuem para o fenómeno de tropismo celular e tipo de infeção, o objetivo seguinte foi descobrir a função molecular de proteínas Inc específicas. Usando ensaios de "yeast-two hybrid", o objetivo foi o de descobrir parceiros de interação entre uma biblioteca de proteínas humanas e as proteínas Inc CT228, CT249 e CT288. De entre os resultados obtidos, caracterizámos a interação entre CT288 e a proteína humana "coiled coil domain containing protein 146" (CCDC146). CCDC146 tem uma função desconhecida, mas sabe-se que localiza no centrossoma. Usámos diferentes técnicas bioquímicas para validar a interação entre a Inc CT288 e a proteína CCDC146, e experiências de imunofluorescência em células infetadas por *C. trachomatis* para estudar a localização das duas proteínas. Descobrimos que em células infetadas por *Chlamydia*, a CT288 e a CCDC146 co-localizam na membrana do inclusão. Sabe-se que o centrossoma, que funciona como coordenador da arquitetura do citosqueleto da célula eucariótica, é manipulado por *C. trachomatis* durante todo o seu ciclo infecioso. Além disso, já foram descritas outras interações entre proteínas centrossomais da célula hospedeira e proteínas Inc. Assim, formulamos a hipótese de que a interação entre a Inc CT288 e a proteína centrossomal CCDC146 pode contribuir para a dinâmica que existe entre o inclusão e o centrossoma.

Em conclusão, nesta tese aumentámos o conhecimento acerca da patogénese molecular de infeções por *C. trachomatis*, particularmente no fenómeno do tropismo celular e tipo de infeção. Descobrimos também um novo alvo eucariota de uma proteína Inc.

Dissertation Outline

This thesis is divided in five chapters. Chapter 1 gives a general picture of infections by *Chlamydia*, and in more detail, by *Chlamydia trachomatis*. It contains also a description of the known *Chlamydia*-host interactions, focusing on the chlamydial developmental cycle and on the type III secretion system (T3SS), along with a brief comparison with other intravacuolar pathogens. Data presented on Chapter 2 and 3 consist of published material supplemented with relevant unpublished results. Chapter 2 describes the identification of type III secretion (T3S) signals in chlamydial inclusion membrane (Inc) proteins using *Yersinia enterocolitica* as heterologous host bacteria. T3S signals were identified in the majority of the Inc proteins, suggesting that they are translocated into the host cell by this transport mechanism. Chapter 3 describes phylogenetic, molecular evolution, and gene expression analyses aimed to determine if Inc proteins could play a role in the type of infection and tissue tropism associated with *C. trachomatis* strains. The results show that a subgroup of Inc proteins could be involved in the distinct tropism and invasiveness associated with *C. trachomatis* LGV strains. Chapter 4 describes data that are part of a manuscript in preparation, and was aimed at finding host cell interacting partners for Incs. A number of candidate host cell interacting partners for Incs CT228, CT249, and CT288 were found, however, during the course of this thesis it was only possible to validate the interaction between CT288 and the host cell centrosomal protein CCDC146. Chapter 5 contains a general discussion about the importance of the results obtained and of the possible future lines of research.

Abbreviations

ARF1	Adenosine diphosphate (ADP)-ribosylation factor 1
ATCC	American Type Culture Collection
ATP	Adenosine triphosphate
BHI	Brain heart infusion
BLAST	Basic local alignment search tool
cDNA	Complementary DNA
COG	Conserved oligomeric Golgi
co-IP	Co-immunoprecipitation
CPAF	Chlamydial protease-like activity factor
DMEM	Dulbecco's modified eagle medium
DNA	Deoxyribonucleic acid
ECDC	European Centre for Disease Prevention and Control
EDTA	Ethylenediamine tetraacetic acid
EGFP	Enhanced green fluorescent protein
ERK	Extracellular signal-regulated kinases
FBS	Fetal bovine serum
GFP	Green fluorescent protein
GTPase	Guanosine triphosphate hydrolases
HA	Human influenza hemagglutinin epitope
HBSS	Hank's balanced salt solution
IκBα	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha
LB	Lysogeny broth
LPS	Lipopolysaccharide
MAPK	Mitogen-activated protein kinases
MOI	Multiplicity of infection
mRNA	Messenger RNA
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NGS	Next-generation sequencing
NOD1	Nucleotide-binding oligomerization domain-containing protein 1
OD_{600}	Optical density at 600 nm
ORF	Open reading frame
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PMSF	Phenylmethylsulfonyl fluoride
RNA	Ribonucleic acid
rpm	Rotations per minute
rRNA	Ribosomal RNA
RT-PCR	Reverse transcription coupled with PCR
RT-qPCR	Real-time quantitative PCR
SDS	Sodium dodecyl sulfate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
SEM	Standard error of the mean
siRNA	Small interfering RNA

SNARE	Soluble N-ethylmaleimide-sensitive factor attachment protein receptor
sRNA	Small RNA
TCA	Trichloroacetic acid
TGN	*Trans*-Golgi network
TLR	Toll-like receptor
USA	United States of America
UV	Ultraviolet
VAMP	Vesicle associated membrane protein
WHO	World Health Organization

List of Figures

List of Tables

Table of Contents

CHAPTER 1 - General Introduction

*T**he author of this dissertation wrote the whole Chapter based on the referred bibliography.***

1.1 The biology of *Chlamydia*

1.1.1 Diversity

Members of the phylum *Chlamydiae* are closely related Gram-negative bacteria, biologically characterized by obligate intracellular growth within eukaryotic cells. The diversity of *Chlamydiae*, their host range, and their occurrence in the environment is a hallmark, as it includes bacterial species that infect or establish symbiotic relationships with different vertebrates, invertebrates, and even unicellular eukaryotes such as amoeba (Everett *et al.*, 1999; Horn, 2008).

Currently, *Chlamydiae* is divided in eight families: *Chlamydiaceae*, *Parachlamydiaceae*, *Criblamydiaceae*, *Waddliaceae*, *Rhabdochlamydiaceae*, *Simkaniaceae*, *Clavochlamydiaceae,* and *Piscichlamydiaceae*. The *Chlamydiaceae* family includes the most important pathogens of humans and other animals, while the other seven families include a growing number of *Chlamydia*-like organisms that are mainly symbionts of eukaryotes (Everett *et al.*, 1999; Horn, 2008).

The *Chlamydiaceae* have been divided in two genera (*Chlamydia* and *Chlamydophila*) (Everett *et al.*, 1999; Horn, 2008). However, since the proposal of this division, a lot of discussion was raised concerning whether the distinction between *Chlamydia* and *Chlamydophila* was necessary. In addition, the majority of the *Chlamydia* research community did not follow this nomenclature. This led to the recently recommended emended descriptions of the family *Chlamydiaceae* and the proposal of a single genus to include all currently recognized species (Sachse *et al.*, 2015). By this new definition, the family *Chlamydiaceae* is composed of the genus *Chlamydia* encompassing 11 described species and one recently proposed new species (Vorimore *et al.*, 2013; Sachse *et al.*, 2014; Sachse *et al.*, 2015):

- *Chlamydia trachomatis* is a human-specific pathogen that causes ocular and sexually transmitted genital infections, and is considered a major health problem worldwide (Wright *et al.*, 2007; Bébéar and de Barbeyrac, 2009).
- *C. pneumoniae* is another human pathogen, and is one of the most prevalent causes of respiratory infections (Blasi *et al.*, 2009; Kern *et al.*, 2009). It accounts for ~15% of community-acquired pneumonia and is associated with exacerbations of chronic bronchitis, pharyngitis and asthma (Blasi *et al.*, 2009; Kern *et al.*, 2009). *C. pneumoniae* has also been linked with other diseases such as atherosclerosis, central nervous system disorders, and Alzheimer's disease, and there are reports suggesting that *C. pneumoniae* non-human infections can lead to respiratory and vascular pathologies in a broad spectrum of animals and reptiles (Bodetti *et al.*, 2002; Roulis *et al.*, 2013).
- *C. psittaci* has also a great impact in the veterinary economy and infects approximately 450 different bird species, including psittacine birds, pigeons, ducks, geese, and turkeys (Knittler *et al.*, 2014). In birds, *C. psittaci* is responsible for psittacosis, one of the most significant animal diseases, manifested most frequently by pneumonia, conjunctivitis and enteritis, among other pathologies. *C. psittaci* is considered an important zoonotic agent as transmission to humans often occurs, causing severe respiratory diseases and even death (Knittler *et al.*, 2014).
- *C. abortus* is of significant economic importance because it can colonize the placenta of sheep and goats, leading to abortion in these animals. It is also responsible for zoonotic infections and is considered a major risk factor for abortion in pregnant women who are exposed (Longbottom and Coulter, 2003).
- *C. muridarum* is a natural mouse pathogen that causes pneumonitis (Ramsey *et al.*, 2009; De Clercq *et al.*, 2013).

- *C. pecorum* is an important pathogen of domesticated livestock including cattle, sheep, goats, and pigs (Reinhold *et al.*, 2011; Bachmann *et al.*, 2014). It is responsible for a wide range of diseases such as polyarthritis, pneumonia, conjunctivitis, and encephalomyelitis, and has been associated with gastrointestinal and urogenital tract diseases, and abortion. This pathogen is also known to cause ocular and genital infections in koalas, and is considered a major contributing factor to the decline of the population of these animals in Australia (Polkinghorne *et al.*, 2013).
- *C. felis* is a common agent of conjunctivitis and upper respiratory tract diseases in cats although, there are also reports of humans being infected with this bacterium (Cai *et al.*, 2002).
- *C. caviae* causes conjunctivitis and genital tract infections in guinea pigs (Read *et al.*, 2003; Neuendorf *et al.*, 2015).
- *C. suis* infects mainly pigs and is associated with porcine conjunctivitis, pneumonia, and various reproductive disorders (Schautteet and Vanrompay, 2011; Donati *et al.*, 2014).
- *C. avium* strains can infect pigeons and members of psittacine birds. Infections appear to be asymptomatic and widely disseminated, still, little is known about whether can infect other animals (Sachse *et al.*, 2014).
- *C. gallinacea* strains were recovered from samples of chickens, guinea fowls, and turkeys. As *C. avium*, infections are asymptomatic and widely disseminated (Sachse *et al.*, 2014).

The newly proposed species is *C. ibidis*, and was recovered from feral African Sacred Ibises (Vorimore *et al.*, 2013). Only a few strains were collected and so far there is no evidence of a pathogenic potential. They appear to be innocuous in avian species but the health risk to humans or other animals is still unknown (Vorimore *et al.*, 2013).

Regarding the *Chlamydia*-like organisms, there are reports suggesting an association between *Parachlamydia acanthamoebae*, *Simkania negevensis*, and *Waddlia chondrophila* in human respiratory diseases and cases of abortion in women (Horn, 2008; Lamoth *et al.*, 2015). However, these associations were mostly based on serological and molecular data and strains were never isolated from a patient and neither their presence was demonstrated at the site of infection (Horn, 2008; Lamoth *et al.*, 2015).

1.1.2 Genomics of Chlamydia

Until recently, the major limitations in the development of a host-free growth system and in the development of genetic methodologies to manipulate *Chlamydia* hampered the understanding of the biology of these organisms. However, about 15 years ago, with the beginning of the genomics era, a revolution started with the first sequenced genome of a *C. trachomatis* strain (Stephens *et al.*, 1998). Subsequent advances in the NGS technology allowed the increase of the number of available *Chlamydia* genomes to more than 100, encompassing almost all recognized species.

Comparative analyses of the genomes of four environmental *Chlamydia*-like species and a large number of *Chlamydia* species indicated that the genomes of these two distinct groups of bacteria are significantly divergent, and that the genomes of *Chlamydia*-like species are also divergent within this group. In particular, the genomes of *Chlamydia*-like bacteria display little or no synteny neither to genomes of *Chlamydia* nor to each other, and they are two to threefold larger than those of *Chlamydia* species (Griffiths *et al.*, 2006; Collingro *et al.*, 2011). In contrast, although some differences are found between the genomes of the various *Chlamydia* species, the majority of the core genes and

pathways are conserved, and a high degree of synteny is maintained (Griffiths *et al.*, 2006; Harris *et al.*, 2012; Knittler *et al.*, 2014). The typical *Chlamydia* chromosome consists of approximately 1.0 to 1.2 Mb and 1000-1200 protein-coding genes, depending on the species. Analyses of the sequenced genomes revealed that *Chlamydia* have the minimal machinery required for DNA replication, transcription and translation, for delivery of proteins by a type III secretion system, for basic lipid metabolism, and for essential functions in aerobic respiration (Stephens *et al.*, 1998; Yao *et al.*, 2015). The relatively small genome size of *Chlamydia* species, when compared to other bacteria with small genomes, suggests that during adaptation to its host they lost a large number of genes, most probably because the environment where they thrive no longer required those functions (Mendonça *et al.*, 2011). In fact, *Chlamydia* have several incomplete metabolic pathways, such as: the tricarboxylic acid cycle, because genes for citrate synthase, aconitase, and isocitrate dehydrogenase are not present in the genome; and for biosynthesis of all amino acids, as for example, phenylalanine and tyrosine, where the enzymes responsible for the last step of biosynthesis of these amino acids are present in the genome, but enzymes for the preceding steps are not (Stephens *et al.*, 1998). Therefore, *Chlamydia* must rely on their eukaryotic hosts for obtaining those nutrients.

The differences in the genomes of *Chlamydia* are mainly restricted to a region of 10 to 50 kb, designated as the plasticity zone (PZ). This region contains several genes that might encode virulence factors, including the *Chlamydia* cytotoxin and putative membrane attack complex/perforin (MACPF), and phospholipase D enzymes. The PZ is often a focus in comparative genomics analyses of *Chlamydia*, as the presence or absence of these genes could play a role in host specificity (Harris *et al.*, 2012; Knittler *et al.*, 2014). Other genes that show pronounced genomic

differences include: the genes encoding for the highly variable polymorphic membrane proteins, a protein family unique to *Chlamydia* (Tan *et al.*, 2006); the gene encoding the type III secretion system translocated actin recruiting phosphoprotein (Tarp) (Voigt *et al.*, 2012); and genes of the biotine and pyrimidine pathways (Voigt *et al.*, 2012).

An additional feature of most *Chlamydia* species is the presence of a highly conserved plasmid of about 8 kb and encoding eight ORFs (Carlson *et al.*, 2008; Gong *et al.*, 2013). This plasmid appears to have an important role during infection and there is also a strong selection mechanism to maintain it (Carlson *et al.*, 2008; Gong *et al.*, 2013). Evolutionary studies revealed that the plasmid co-evolved with the chromosome, as it exhibits segregation of species/strains that co-relate with their host range and tissue tropism (Read *et al.*, 2003; Harris *et al.*, 2012; Ferreira *et al.*, 2013). Plasmid-free *Chlamydia* strains rarely occur in nature and studies using these strains indicated that the plasmid is important for glycogen synthesis, in modulating the host cell inflammatory response, and in the regulation of expression of a set of chromosomal genes (O'Connell *et al.*, 2007; Wang *et al.*, 2011; Song *et al.*, 2013; Porcella *et al.*, 2015).

Another important point on the chlamydial biology research was the discovery that recombination often occurs in *Chlamydia*. It is known that these organisms possess a complete system for DNA recombination encoded in the genome and there are several reports showing that recombination contributed significantly for the genetic diversity of *Chlamydia,* including incorporation of foreign DNA (DeMars *et al.*, 2007; Binet and Maurelli, 2009; Jeffrey *et al.*, 2010; Joseph *et al.*, 2011; Ferreira *et al.*, 2012; Harris *et al.*, 2012; Read *et al.*, 2013).

1.1.3 Transcriptomics and proteomics of Chlamydia

With the data generated from the first sequenced genomes, the attention then moved towards the products of the genes. Several proteomic studies were developed and aimed at the identification of the complete battery of proteins of *Chlamydia* (Vandahl *et al.*, 2001; Skipp *et al.*, 2005; Saka *et al.*, 2011). Similar assays were done focusing on the host proteome, to evaluate how infection with *Chlamydia* alters host protein stability (Olive *et al.*, 2014). Transcriptomic studies were also performed to assess the changes of expression of chlamydial and host genes during infection (E. I. Shaw *et al.*, 2000; Belland *et al.*, 2003; Nicholson *et al.*, 2003; Borges *et al.*, 2010; Albrecht *et al.*, 2011; Ferreira *et al.*, 2013; Omsland *et al.*, 2013; Porcella *et al.*, 2015). These studies revealed that *Chlamydia* are packed with proteins required for central metabolism, energy utilization, and in the delivery of virulence proteins in distinct phases of the developmental cycle.

1.1.4 New advances in the chlamydial biology

In the last five years, we witnessed a breakthrough in the *Chlamydia* research field. In 2011, the report of a shuttle vector and a stable system for genetic transformation of *C. trachomatis* allowed for the first time the analysis of gene function by complementation (Wang *et al.*, 2011). Since then, several adaptations of this system were developed with the design of inducible systems that incorporate different fluorescent proteins (Agaisse and Derré, 2013; Gérard *et al.*, 2013; Wickstrum *et al.*, 2013). Novel insights into the molecular function of chlamydial proteins were already achieved using this methodology (Agaisse and Derré, 2014; Bauler and Hackstadt, 2014; Mirrashidi *et al.*, 2015; Mueller and Fields, 2015). Another method recently implemented to genetically manipulate *Chlamydia* is the generation of null mutants by chemical mutagenesis (Kari *et al.*, 2011; Nguyen and Valdivia, 2012; Kokes *et al.*, 2015).

Although with this method libraries of mutants can be constructed, it generates random mutations in the genome and still requires a subsequent step of whole genome sequencing for validation. In addition, the mobile group II intron system and other systems have been optimized to perform targeted inactivation of chromosomal genes in *Chlamydia* (Johnson and Fisher, 2013; Thompson *et al.*, 2015). Very recently, a novel system for targeting chlamydial genes for deletion or allelic exchange was developed for *C. trachomatis* (Mueller *et al.*, 2016). In this system, a novel suicide vector allows the generation of fully genetically tractable *C. trachomatis* and the process can be easily monitored by the use of fluorescence reporters (Mueller *et al.*, 2016).

1.1.5 The "chlamydial anomaly" resolved

For a long time, the existence or not of peptidoglycan in *Chlamydia* has generated a lot of controversy. Conflicting data between genomics and antibiotic resistance, which indicated that peptidoglycan exists in *Chlamydia* and the unsuccessful attempts in detection and purification of chlamydial peptidoglycan components led to the so-called "chlamydial anomaly" (Chopra *et al.*, 1998). However, recent advances in mass spectrometry and in new cell-wall labeling methods enabled the structural characterization of peptidoglycan components from *C. trachomatis*, and to solve a 50 year mystery (Liechti *et al.*, 2014; Packiam *et al.*, 2015).

1.1.6 Tissue culture and animal models

As *Chlamydia* organisms cannot grow on conventional bacteriological medium, tissue culture models have been instrumental to study the biology of *Chlamydia*. Cultured cell lines allowed researchers to isolate, propagate, and purify large quantities of chlamydial organisms, and have

made possible advances in the chlamydial biology through further proteomic, microscopy and other biochemical experiments (Scidmore, 2005).

Animal models are indispensable to study bacterial infections. The most commonly used animal models to study female genital tract infections with *Chlamydia* are the mouse, the guinea pig, the pig, nonhuman primate models, and to a less extent, the rat and the rabbit (De Clercq *et al.*, 2013; Neuendorf *et al.*, 2015). The mouse is used mainly with *C. trachomatis* and *C. muridarum*, the pig and nonhuman primates are mainly used with *C. trachomatis*, and the guinea pig is used with *C. caviae*. With these models it is possible to mimic several aspects of the chlamydial infection, pathogenesis, immunity and treatment, such as the specific tissue tropism, sexual transmission, disease manifestations, and vaccine development (De Clercq *et al.*, 2013; Neuendorf *et al.*, 2015).

1.1.7 Antimicrobial resistance

Resistance to antimicrobials in *Chlamydia* remains rare. Chlamydial organisms are susceptible to most common antimicrobials; however, *in vitro* tests of antimicrobial susceptibility are challenging because the results are variable, and depend on the cell line used and at which time the antimicrobial is added after infection (Knittler *et al.*, 2014; Kong and Hocking, 2015). Therapeutic failures have been attributed to events of heterotypic resistance, where a small proportion of resistant organisms have an advantage over a mostly susceptible population (Somani *et al.*, 2000). This is in agreement with reports that demonstrate that *Chlamydia* easily and rapidly develop resistance when bacteria are continually exposed to antimicrobials *in vitro* (Binet and Maurelli, 2005; DeMars *et al.*, 2007). Due to the characteristic intracellular environment where *Chlamydia* thrive, the opportunity for the acquisition of

antimicrobial resistance genes from other organisms is limited. Instead, the acquisition of antimicrobial resistance in *Chlamydia* occurs mainly through point mutations that result in the overexpression or alteration of the inhibitory target (Binet and Maurelli, 2005).

1.1.8 Vaccine development

The development of vaccines is of great importance due to the increasing rates of mainly asymptomatic *C. trachomatis* infections and the adverse long-term consequences resulting from these infections. However, the lack of knowledge of the genital tract and of the protective immune responses triggered by *Chlamydia* prevents further progress. A vaccine against ovine abortions caused by *C. abortus* was already implemented with success and a vaccine currently under development to control chlamydial infections in the koala is showing promising results, indicating that vaccination against *C. trachomatis* is also possible in humans (Polkinghorne *et al.*, 2013). Several attempts for the development of a human vaccine have been made but the results were ambiguous. In most cases, there were reports of enhanced disease after immunization (Stary *et al.*, 2015). Recently, preliminary experiments performed on humanized mice with a vaccine made of a conjugation of nanoparticles with UV light-inactivated *C. trachomatis* showed that the immune response generated by this type of vaccine elicited protection in the mice against the challenge with *C. trachomatis* strains (Stary *et al.*, 2015).

1.2 *Chlamydia trachomatis*

1.2.1 C. trachomatis *strains and infectious diseases*

C. trachomatis strains are divided in two biovars: trachoma and lymphogranuloma venereum (LGV) (Schachter, 1999). Differential immunoreactivity of the major outer membrane protein (MOMP) of *C. trachomatis*, or genotyping by PCR and sequencing of *ompA* (the gene encoding MOMP), allows the classification of the strains in 15 major serovars (A-L) (Schachter, 1999). The trachoma biovar comprises strains that mainly cause localized infections of the epithelial surface of the conjunctiva (serovars A-C) and genital mucosa (serovars D-K). Infection of conjunctival epithelial cells lead to conjunctivitis, with severe inflammation and scarring of the conjunctiva, ultimately resulting in trachoma, the most frequent cause of preventable blindness (Wright *et al.*, 2007). Trachoma is endemic in 51 countries and is responsible for the visual impairment of about 2.2 million people, of whom 1.2 million are irreversibly blind (WHO, 2014). Infection of genital epithelial cells can progress to pelvic inflammatory disease in women, and urethritis in men. As most of these genital infections are asymptomatic, there is a high probability for the development of salpingitis and epididymitis, and, especially in women, tubal factor infertility and ectopic pregnancy (Bébéar and de Barbeyrac, 2009). The rate of cases of *C. trachomatis* sexually transmitted diseases reported to the ECDC in Europe in 2012 was 184 per 100000 population (ECDC, 2014), and the WHO reported for the USA in 2008 more that 100 million new cases of *C. trachomatis* sexually transmitted infections (WHO, 2008). Strains of the serovars L1-L3 are responsible for a more invasive urogenital disease, LGV, due to the ability of these strains to infect macrophages and spread to regional lymph nodes (Bébéar and de Barbeyrac, 2009). In 2012, 830 cases of LGV were reported in eight countries of Europe, a number that has been increasing over the years (ECDC, 2014). Portugal lacks an efficient

screening program for *C. trachomatis* infections, but since 2014 it is mandatory to report cases of infection with this bacterium. Still, a study based on data collected in one diagnostics laboratory reported 2817 (7.7%) positive *C. trachomatis* cases among 36621 individuals tested (Dinis *et al.*, 2015).

As mentioned above in section 1.1.7, antimicrobial resistance in *C. trachomatis* rarely occurs, and once diagnosed, infections are easily and effectively treated with regular doses of azithromycin or doxycycline. However, once a severe infection and pathology are established, treatment is often less effective and there is a risk of serious reproductive damage, including a link with cervical cancer (Kong and Hocking, 2015). The common etiological agent of cervical cancer is the human papilloma virus (HPV); however, early epidemiological studies have shown that sexual activity is also a risk factor for cervical carcinoma (Koskela *et al.*, 2000; Dahlström *et al.*, 2011). This led to the suggestion that sexually transmitted diseases, particularly exposure to *C. trachomatis*, could contribute to the development of identical pathologies. *In vitro* studies demonstrated that infection with *C. trachomatis* leads to defects in centrosome dynamics and in DNA repair mechanisms that result in abnormal chromosomal segregation to daughter cells during cell division (Grieshaber *et al.*, 2006; Chumduri *et al.*, 2013). Chromosome instability is known to be one of the main factors involved in tumor development, and therefore it is possible that *C. trachomatis* genital infections could contribute to cervical cancer (Weitzman and Weitzman, 2014; Nam *et al.*, 2015).

1.2.2 Tissue tropism and type of infection

The rapid expansion in the number of available sequenced genomes of *C. trachomatis* brought new insights into the differences in type of

infection (invasive or non-invasive) and tissue tropism (eyes, genitals, and lymph nodes) observed with *C. trachomatis* strains. Interestingly, analyses of more than 50 sequenced genomes showed that ocular (serovars A-C), urogenital (serovars D-K), and LGV (serovars L1-L3) strains exhibit >98% identity and a high degree of synteny (Stephens *et al.*, 1998; Carlson *et al.*, 2005; Thomson *et al.*, 2008; Seth-Smith *et al.*, 2009; Jeffrey *et al.*, 2010; Unemo *et al.*, 2010; Somboonna *et al.*, 2011; Harris *et al.*, 2012).

The differences in type of infection and tissue tropism among *C. trachomatis* strains must reside in the little dissimilarity in their genetic background, and there are a number of *C. trachomatis* genes that have been associated with this:

- The clearer example is the distinction between ocular strains from urogenital and LGV strains through specific mutations in the tryptophan (*trpBA*) operon (Allan C. Shaw *et al.*, 2000; Fehlner-Gardiner *et al.*, 2002; Caldwell *et al.*, 2003). In ocular strains, polymorphisms in the tryptophan synthase gene B (*trpB*) result in a nonfunctional protein. In urogenital and LGV strains there is a strong selective pressure to maintain a functional tryptophan synthase, as this mechanism is necessary for the bacteria to produce tryptophan in response to the host induction of tryptophan-degrading enzymes (Allan C. Shaw *et al.*, 2000; Fehlner-Gardiner *et al.*, 2002; Caldwell *et al.*, 2003).
- Another genetic difference that can explain the tissue tropism of *C. trachomatis* strains is the variation in the *Chlamydia* cytotoxin gene that is located in the PZ (Carlson *et al.*, 2004). This gene has extensive mutations and deletions among the different serovars that clearly distinguish the three groups of strains. The urogenital strains all encode an intact N-terminal-encoding toxin that includes both the UDP-glucose binding and

glycosyltransferase domains, the LGV strains lack both of these domains, while the ocular strains only encode the UDP-glucose binding domain (Carlson *et al.*, 2004). It has been hypothesized that the N-terminal region of the cytotoxin could play a key role in allowing *C. trachomatis* to colonize the genital tract.

- In the PZ there is also a group of genes encoding for proteins with identity to the phospholipase D (PLD), that might be involved in the acquisition and processing of host lipids (Nelson *et al.*, 2006). In the genomes of *C. trachomatis* there is an accumulation of multiple deletions and frameshift mutations in these genes that may be associated with the strain-specific features of infection (Nelson *et al.*, 2006).
- Polymorphic membrane proteins (Pmps) participate in adhesion of the bacterium to the host, and also present characteristics that might contribute to tissue tropism (Becker and Hegemann, 2014; Kari *et al.*, 2014). Phylogenetic reconstructions based in the amino acid sequence of Pmps showed a separation of the strains into groups with similar disease properties (Gomes *et al.*, 2006).
- Tarp is a protein delivered by *Chlamydia* into host cells that participates in the invasion process by inducing actin polymerization at the site of bacterial entry (Jewett *et al.*, 2010). Nucleotide and amino acid sequence analysis of the gene encoding Tarp and of Tarp itself, respectively, indicated an evolutionary relationship with disease phenotype, separating LGV and ocular strains from urogenital strains. Tarp proteins from LGV strains contain higher number of tyrosine-rich repeat regions and fewer predicted actin binding domains, in contrast with ocular strains that contain more actin binding domains and fewer tyrosine-rich repeat regions (Lutter *et al.*, 2010).

It has also been suggested that the plasmid plays a key role in tissue tropism. A transcriptomics study found for the majority of the plasmid ORFs lower expression profiles for ocular strains (Ferreira *et al.*, 2013). Further genome comparisons and single nucleotide polymorphism analyses highlighted additional spots putatively involved in the type of infection and tissue tropism (Brunelle *et al.*, 2004; Nunes *et al.*, 2008; Thomson *et al.*, 2008; Joseph *et al.*, 2011; Borges *et al.*, 2012; Lutter *et al.*, 2012). For example, LGV strains appear to have lost some metabolic capacity, as shown by pseudogenes for fumarate hydratase and succinate dehydrogenase.

1.3 *Chlamydia* developmental cycle

1.3.1 Elementary bodies and reticulate bodies

All *Chlamydiae* have a unique infectious cycle, different from other intracellular bacteria in that growth is associated with a biphasic developmental cycle, alternating between two morphologically distinct forms: the elementary body (EB) and the reticulate body (RB) (Figure 1.1) (AbdelRahman and Belland, 2005). EBs are the infectious and non-replicative form, and are responsible for the dissemination of the infection. They are relatively small (approximately 0.3 μm of diameter) and have a highly cross-linked outer membrane that confers protection in the outside environment. In contact with a host cell, EBs present a polarized architecture, with a highly compact DNA material on one side and an array of type III secretion systems on the other side, oriented to the host cell (AbdelRahman and Belland, 2005; Nans *et al.*, 2014). Upon infection, EBs are internalized in a membrane-bound vacuole, where they start a primary differentiation into RBs (Figure 1.1), the non-infectious and replicative form, which are larger than EBs (approximately 1 μm of diameter) and have a high metabolic activity (AbdelRahman and Belland, 2005). After repeated cycles of bacterial binary fission, the vacuole enlarges and becomes filled with progeny. Strictly speaking, it is the larger vacuole that corresponds to an “inclusion”, but this term is normally used to refer to chlamydial vacuole since its formation soon after bacterial invasion. At a certain stage of *Chlamydia* cell division, a second differentiation of the bacteria occurs, back to EBs (Figure 1.1). At the end of the cycle, EBs are released, either by host cell lysis or extrusion, and are able to start a new cycle.

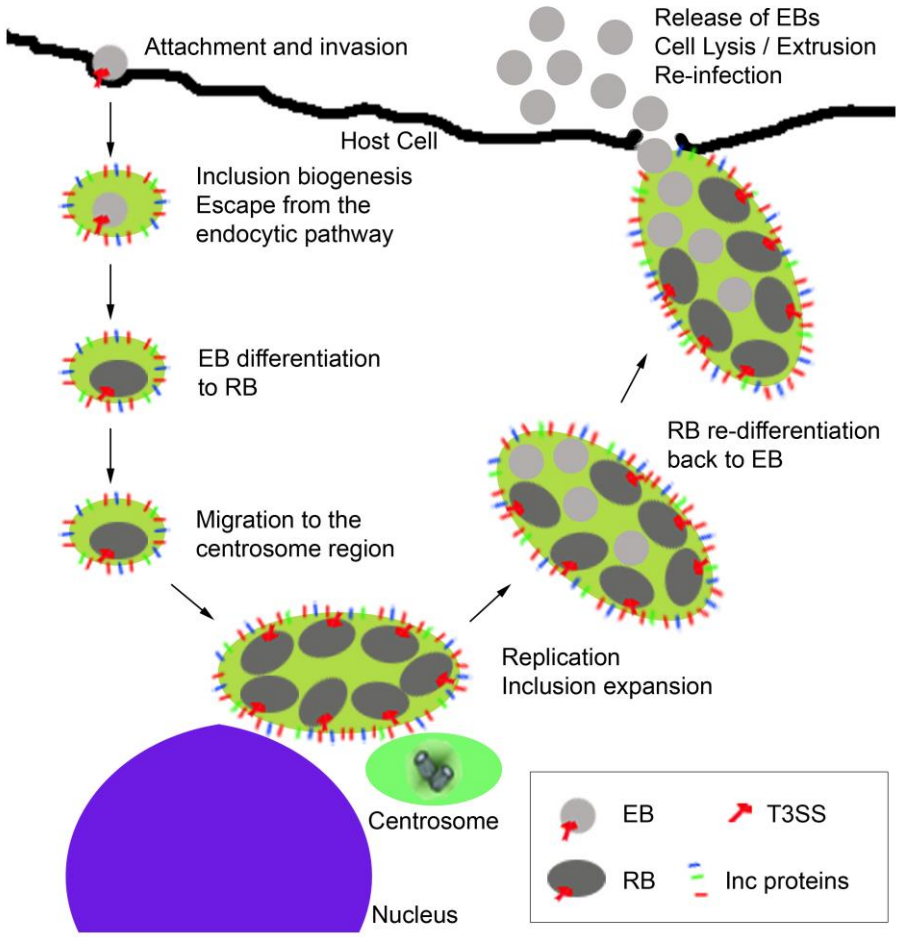


Figure 1.1 – Representation of the *Chlamydia* biphasic developmental cycle. Chlamydial organisms cycle between two different morphological forms: the EBs (light grey), the infectious form, and the RBs (dark grey), the replicative form. EBs invade the host cell and localize within a membrane-bound vacuole, known as inclusion. Within this vacuole, EBs differentiate into RBs. The inclusion then migrates towards the centrosome (the main microtubule organizing center [MTOC]), where RBs start to replicate. As the inclusion expands with progeny, RBs re-differentiate back to EBs that will exit the cell by extrusion or cell lysis. *Chlamydia* use a type III secretion system (T3SS) to deliver effector proteins into host cells (see section 1.4). These effectors manipulate host processes at different stages of the developmental cycle. Some effector proteins, known as inclusion membrane (Inc) proteins, localize at the inclusion membrane. Adapted from (AbdelRahman and Belland, 2005).

Under certain conditions of stress, the developmental cycle is deviated from the normal route, resulting in the appearance of an aberrant form (often called AB) and the bacteria remain in a state of persistence (AbdelRahman and Belland, 2005; Schoborg, 2011). Persistence in *Chlamydia* is defined as an *in vitro* reversible condition in which the bacteria remain in a viable but culture-negative state. The ABs are normally described as enlarged RBs that neither differentiate back to EBs nor replicate normally. It is unknown, however, whether this persistent form occurs also *in vivo* (Schoborg, 2011).

1.3.2 Regulation of the chlamydial developmental cycle

The precise mechanisms that regulate the transition between *Chlamydia* EBs and RBs unknown. Soon after internalization, *Chlamydia* start using its internal machinery for protein synthesis, and this event is consistent in time with the first differentiation step. Then, another stage of protein synthesis is initiated to allow bacterial growth and division. The last period of protein synthesis is triggered at the same time that RBs re-differentiate back to EBs. The developmental cycle takes about 48 to 72 h, depending on the species (E. I. Shaw *et al.*, 2000; Vandahl *et al.*, 2001; Belland *et al.*, 2003; Nicholson *et al.*, 2003; Skipp *et al.*, 2005; Saka *et al.*, 2011; Omsland *et al.*, 2013). Some authors suggested the presence of a quality control system that regulates the chlamydial transition from RBs to EBs. In the presence of some inhibitors, RBs seem to fail the expression of selective late-cycle genes and the differentiation to EBs is interrupted (Nguyen *et al.*, 2011). The highly condensed DNA exhibited by EBs has been proposed to confer a complete transcriptional shutdown. However, recent studies indicated that despite the compact chromosome, EBs are still able to perform basic metabolic activities (Omsland *et al.*, 2013; Sixt *et al.*, 2013).

It is unknown how specific genes are selectively transcribed at each stage of the developmental cycle. Three σ subunits of RNA polymerase have been identified in *Chlamydia*: σ^{66} is the homolog of *Escherichia coli* σ^{70}, and two alternative σ subunits show sequence identity to *E. coli* σ^{54} and *Bacillus subtilis* σ^{28}. Early- (2 to 8 h), mid- (12 to 20 h), and late-cycle (30 to 48 h) chlamydial genes are mostly transcribed by the σ^{66} RNA polymerase (Tan *et al.*, 1998; Mathews and Timms, 2006; Hefty and Stephens, 2007; Mallios *et al.*, 2009; Case *et al.*, 2010). This implies that there must be specific mechanisms that differentially regulate the temporal classes of gene expression. Expression of mid-cycle genes has been proposed to be regulated by differential levels of DNA supercoiling in mid-cycle (Niehus *et al.*, 2008; Cheng and Tan, 2012). Expression of late chlamydial genes has been shown to be mediated also by the σ^{28} (Mathews and Timms, 2006). The chlamydial protein early upstream ORF (EUO) has been proposed as a repressor of late genes, as it is encoded in all *Chlamydia* genomes, is expressed early in the cycle, and contains a DNA-binding domain described to selectively bind to promoter regions of known chlamydial late genes (Rosario and Tan, 2012). The chlamydial protein general regulator of genes A (GrgA) was also described to be a DNA-binding protein that can stimulate transcription from a range of σ^{66}-dependent promoters, suggesting its role as a regulator of σ^{66}-dependent transcription in *Chlamydia* (Bao *et al.*, 2012).

1.3.3 Manipulation of host cell processes at different stages of the chlamydial developmental cycle

In all stages of the developmental cycle, *Chlamydia* release effector proteins into the host cell cytosol by a type III secretion system, a common virulence mechanism used by Gram-negative bacteria (described in detail in section 1.4). In general, the effector proteins are

responsible for the manipulation of host cell processes to the benefit of the bacteria (Mota and Cornelis, 2005; Galán *et al.*, 2014).

1.3.3.1 Adhesion

Infection by *Chlamydia* primary starts with adhesion of EBs to the host cell plasma membrane. This process is mediated by electrostatic interactions of the bacteria and binding to receptors. Different bacterial adhesins have been proposed, such as MOMP and Pmps (Kari *et al.*, 2014; Nans *et al.*, 2015). Pmps were described to mediate adhesion of the bacteria to mammalian cells *in vitro* and to be involved in the adaptation of the different chlamydial species to their specific niches (Mölleken *et al.*, 2010; Becker and Hegemann, 2014; Kari *et al.*, 2014). On the host side, among others, heparan sulfate containing glycosaminoglycans, mannose receptors, and Ephrin A2 have been implicated in the adhesion process (Kim *et al.*, 2011; Rosmarin *et al.*, 2012; Karunakaran *et al.*, 2015; Nans *et al.*, 2015; Subbarayal *et al.*, 2015).

1.3.3.2 Invasion

The process of *Chlamydia* entry into host cells is accomplished by bacterial-mediated endocytosis, in which the host cell plasma membrane envelops around the EB, tightly enclosing it until the vacuole is formed (Nans *et al.*, 2015). This process requires delivery of bacterial proteins into host cells (Carabeo *et al.*, 2002; Chin *et al.*, 2012). Tarp and the cytotoxin were described to be delivered immediately after adhesion and to participate directly in the recruitment and remodeling of actin to the sites of attachment (Belland *et al.*, 2001; Carabeo *et al.*, 2007; Lane *et al.*, 2008; Jewett *et al.*, 2010; Thalmann *et al.*, 2010; Bothe *et al.*, 2015). CT694 was also found to be delivered into host cells early in the cycle

and affect the formation of host-cell actin stress fibers (Hower *et al.*, 2009; Bullock *et al.*, 2012). Another protein delivered into host cells by *C. trachomatis* is the translocated early phosphoprotein (TepP), which was reported to act in the entry process by modulating innate immune responses (Chen *et al.*, 2014). A *C. caviae* and *C. pneumoniae* specific protein, *Chla*OTU, was described to be involved in the clearance of ubiquitin at the invasion sites (Furtado *et al.*, 2013).

1.3.3.3 The inclusion

Although *Chlamydia* remain enclosed in the inclusion throughout the entire developmental cycle, they are still able to continually manipulate the host and subvert host pathways to maintain an environment not only suitable for replication and protection from the host but also to scavenge essential nutrients from the host for efficient metabolic activity (Valdivia, 2008; Betts *et al.*, 2009; Kumar and Valdivia, 2009). The inclusion membrane appears to be permeable to small solutes but it is impermeable to components >0.5 kDa. The inclusion membrane is also extensively decorated with inclusion membrane (Inc) proteins, which are thought to play a major role in mediating the cross-talk between the pathogen and the host (Rockey *et al.*, 2002; Mital *et al.*, 2010; Dehoux *et al.*, 2011).

Chlamydia typically replicate within a single large inclusion, containing thousands of bacteria. Even when a cell is infected with more than one individual bacterium, the inclusions fuse into a single vacuole. However, this is not observed for all *Chlamydia* species. In *C. caviae* for example, infection leads to multiple inclusions that do not fuse (Hackstadt *et al.*, 1999; Richards *et al.*, 2013). The fusion of the inclusions was demonstrated to be mediated by an Inc protein, IncA (Hackstadt *et al.*,

1999; Jahn and Scheller, 2006; Delevoye *et al.*, 2008; Ronzone and Paumet, 2013; Ronzone *et al.*, 2014).

1.3.3.4 Interference with host vesicular trafficking pathways

Immediately after internalization, one of the first events promoted by *Chlamydia* is the subversion of the host endocytic pathway (Fields and Hackstadt, 2002; Scidmore *et al.*, 2003). This enables the bacteria to create a unique compartment, devoid of the majority of the markers for early and late endosomes, and lysosomes, thus escaping from one of the major host defense mechanisms (Fields and Hackstadt, 2002; Scidmore *et al.*, 2003). Another essential step for the inclusion biogenesis is the intersection of a different set of vesicles from those involved in lysosomal fusion. Within few hours post-infection, *Chlamydia* intercept the exocytic pathway and acquire endogenous synthesized sphingomyelin, in transit from the Golgi to the plasma membrane (Hackstadt *et al.*, 1996; Scidmore *et al.*, 1996; Elwell *et al.*, 2011).

Infection with *Chlamydia* induce Golgi fragmentation and formation of Golgi ministacks to enhance lipid acquisition and ensure bacterial replication (Heuer *et al.*, 2009). These organisms also hijack the COG complex to redirect the population of Golgi-derived retrograde vesicles to the inclusion (Pokrovskaya *et al.*, 2012). In addition, *Chlamydia* sequester multivesicular bodies and lipid droplets as another source of lipids and cholesterol (Beatty, 2006; Kumar *et al.*, 2006; Beatty, 2008; Cocchiaro *et al.*, 2008; Saka *et al.*, 2015). The retromer plays a central role in the retrieval of several different cargo proteins from the endosome to the TGN. During infection, *Chlamydia* affect the function of the retromer by interacting with sorting nexins (SNXs) (Kabeiseman *et al.*, 2013; Mirrashidi *et al.*, 2015). Disruption of the retromer function is

benefic for the bacteria since depletion of retromer components enhances progeny production (Mirrashidi *et al.*, 2015).

The eukaryotic small GTPase Rab family of proteins regulates host intracellular membrane trafficking processes including vesicle budding, tethering, and fusion. There are several studies demonstrating the recruitment of some of these proteins to the inclusion membrane, suggesting a direct implication in the regulation of the trafficking or fusogenic properties of the inclusion (Rzomp *et al.*, 2003; Rzomp *et al.*, 2006; Cortes *et al.*, 2007; Capmany and Damiani, 2010; Moorhead *et al.*, 2010; Leiva *et al.*, 2013; Sherwood and Roy, 2013).

1.3.3.5 Acquisition of nutrients and lipids by non-vesicular pathways

During the developmental cycle, *Chlamydia* need to acquire many biochemical precursors from the host cell for its growth (Stephens *et al.*, 1998). *Chlamydia* accumulate sugars in the form of glycogen in the inclusion lumen. This process is dependent on the plasmid, since plasmidless strains do not exhibit this phenotype (Carlson *et al.*, 2008; Wang *et al.*, 2011). *Chlamydia* are also able to transport host-derived glycerophospholipids into the inclusion lumen to modify and incorporate them in the chlamydial membrane (Wylie *et al.*, 1997). Peroxisomes are imported into the inclusion and its enzymatic capacity is used for the synthesis of bacterial-specific phospholipids (Boncompain *et al.*, 2014).

The association between the inclusion and the endoplasmic reticulum (ER), and the formation of ER-inclusion membrane contact sites (MCSs) has been extensively studied (Giles and Wyrick, 2008; Derré *et al.*, 2011; Dumoux *et al.*, 2012). Several proteins were described to be enriched at MCSs, including the chlamydial inclusion membrane protein IncD, and the host soluble ceramide transfer protein (CERT), stromal

interaction molecule 1 (STIM1), sphingomyelin synthase 2 (SMS2), vesicle-associated membrane protein-associated protein A (VAPA), and VAPB (Elwell *et al.*, 2011; Agaisse and Derré, 2014; Agaisse and Derré, 2015). It was suggested that MCSs could represent portals for the selective Influx of cellular material into the inclusion such as lipids and sphingomyelin. The interaction with the ER also appears to induce the ER stress response, but this response is quickly dampened by *Chlamydia* to promote host cell survival (Shima *et al.*, 2015).

1.3.3.6 Migration to the centrosomal region

Chlamydia then hijack the host minus-end-directed microtubule motor complex dynein to migrate the inclusion along the microtubules towards the centrosome, the major microtubule organizing center (MTOC) in animal cells (Grieshaber *et al.*, 2003; Richards *et al.*, 2013). The Inc protein CT850 was described to be necessary for a proper localization of the inclusion at the centrosome (Mital *et al.*, 2015). The association between the inclusion and the centrosome is maintained throughout the entire developmental cycle and has several consequences for the host cell (Grieshaber *et al.*, 2003; Grieshaber *et al.*, 2006). A loss of control of the centrosome duplication pathway, a block in host cell cytokinesis, and an impairment of the DNA damage repair mechanism leads to amplified centrosome numbers, increased accumulation of host cell nuclei within cells, and the subsequent formation of multipolar spindles that unevenly distribute chromosomes during cell division (Greene and Zhong, 2003; Grieshaber *et al.*, 2006; Alzhanov *et al.*, 2009; Johnson *et al.*, 2009; Knowlton *et al.*, 2011; Chumduri *et al.*, 2013; Brown *et al.*, 2014; González *et al.*, 2014; Sun *et al.*, 2015).

1.3.3.7 Subversion of the host cytoskeleton

In addition to modifications in the host actin cytoskeleton that are involved in bacterial uptake, *Chlamydia* need to modify the inclusion membrane and alter the cytoskeletal structure of the host cell to allow for the expansion of the inclusion caused by the increasing numbers of bacteria generated during replication (Carabeo *et al.*, 2007; Kumar and Valdivia, 2008; Engström *et al.*, 2015). During infection, the inclusion is encased in a network of host cytoskeletal structures composed of F-actin and intermediate filaments. Disruption of these structures leads to a loss of inclusion shape and leakage of inclusion contents into the host cytoplasm, indicating that this structural scaffold is important for inclusion integrity (Kumar and Valdivia, 2008; Kokes *et al.*, 2015). The chlamydial inclusion protein acting on microtubules (IPAM; CT223) and the inclusion membrane protein for actin assembly (InaC; CT813) were described to be important for the assembly of these superstructures that are crucial to maintain the inclusion shape and vital for the chlamydial developmental cycle (Dumoux *et al.*, 2015; Kokes *et al.*, 2015).

1.3.3.8 Immune detection of **Chlamydia**

Chlamydia infections are detected by the host immune system through the recognition of chlamydial LPS and other pathogen-associated molecular patterns (PAMPs), via TLR, MAPK/ERK, and NOD1 signaling pathways (Bastidas *et al.*, 2013). The NF-κB signaling pathway is immediately shutdown by reversing IκBα ubiquitination through the activity of secreted chlamydial proteins and through the proteolysis of the p65/RelA subunit of NF-κB (Misaghi *et al.*, 2006; Lad, Li, *et al.*, 2007; Lad, Yang, *et al.*, 2007; Bastidas *et al.*, 2013). In addition, *Chlamydia* inhibit apoptosis in the host cell by degrading proteins that control mitochondrial membrane permeabilization, and by interfering with

protein synthesis and upregulating anti-apoptotic proteins (Sharma and Rudel, 2009; Sharma *et al.*, 2011; Flores and Zhong, 2015).

1.3.3.9 Host cell egress

To complete the developmental cycle and ensure proper dissemination and transmission, *Chlamydia* must exit the host cell. Currently there are two proposed mechanisms: cellular lysis and extrusion (Hybiske and Stephens, 2007; Traven and Naderer, 2014). Lysis is a destructive mode of release, consisting of the sequential rupture of the inclusion and cellular membranes by the action of cysteine proteases, leading to host cell death. In contrast, extrusion involves a process in which a portion of the chlamydial inclusion is released, leaving the original cell and the residual inclusion intact (Hybiske and Stephens, 2007). *Chlamydia* encode a protein containing a MACPF domain, used by different parasites to facilitate the egress from cells (Kadota *et al.*, 2004; Rosado *et al.*, 2008; Kafsack *et al.*, 2009; Taylor *et al.*, 2010). *Chlamydia* also recruit elements of the myosin phosphatase pathway to favor the extrusion mechanism via actin polymerization (Lutter *et al.*, 2013).

1.4 Type III secretion in *Chlamydia*

1.4.1 Bacterial protein secretion systems

Bacteria have evolved an array of specialized nanostructures that secrete a wide range of proteins, which participate in several processes of bacterial adaptation and pathogenicity. In Gram-negative bacteria, these protein transport systems are mostly classified as type I secretion system (T1SS), T2SS, T3SS, T4SS, T5SS and T6SS (Costa *et al.*, 2015). The T1, T2, and T5SSs secrete proteins across the two bacterial membranes into the extracellular milieu. The T3, T4, and T6SS are more specialized because they can function as injection devices enabling the delivery of effector proteins from the bacterial cytosol into a eukaryotic or prokaryotic host cell. At a mechanistic level, protein secretion by the T2 and T5SS is a two-step process involving a first translocation step across the inner bacterial membrane that is mediated by the general secretory pathway (Sec system). In contrast, protein transport by the T1, T3, T4, and T6SS occurs in one-step and is Sec-independent (Costa *et al.*, 2015).

The T1SSs are closely related to multidrug efflux pumps and secrete ions, drugs, and toxins. The T2SS secretes mainly hydrolyzing enzymes and toxins to the extracellular milieu. The T4SSs are the most ubiquitous secretion systems in nature, being found in both Gram-negative and Gram-positive bacteria, and in some archaea; their most common role is transfer of DNA between bacteria (conjugation) but they are also required to support the life-cycle of many intracellular pathogens through the delivery of effector proteins into eukaryotic host cells. The T5SS is unique in that the substrate, normally a virulence factor, contains a “auto-transporter” domain that forms a pore in the bacterial outer membrane and drives translocation of a “passenger” domain within the substrate. After translocation, the passenger domain can be cleaved and

released in the extracellular milieu or remain attached to the auto-transporter domain. Because of this, the T5SS is normally known as the auto-transporter system. The T6SSs are known to translocate effector proteins into both prokaryotic and eukaryotic cells but their major role appears to be in bacterial competition, by secreting toxins directly into a neighbor competing bacterium (Costa *et al.*, 2015).

More recently, secretion of curli by Gram-negative bacteria has been termed type VIII secretion (Cao *et al.*, 2014), and proteins involved in gliding motility of Bacteriodetes have been described to be secreted by type IX secretion (McBride and Nakane, 2015). Furthermore, in Gram-positive bacteria (actinobacteria and firmicutes), including pathogens such as *Mycobacterium tuberculosis*, *Staphylococcus aureus* or *Bacillus anthracis*, possess genes encoding a protein secretion machinery that has been termed type VII secretion.

1.4.2 Type III secretion systems

Type III secretion systems (T3SSs) are one of the most complex protein transport systems used by Gram-negative bacteria and have been most intensively studied in pathogenic bacteria such as enteropathogenic and enterohemorrhagic *Escherichia coli* (EPEC and EHEC), *Pseudomonas aeruginosa*, *Salmonella* spp., *Shigella* spp, and *Yersinia* spp. (Mota and Cornelis, 2005; Galán *et al.*, 2014). The assembly of the type III secretion (T3S) apparatus (the injectisome) requires approximately 25 proteins and is composed of a base structure embedded in the bacterial envelope and a needle-like extension that protrudes from the bacterial surface (Figure 1.2). The injectisome is connected to host cell membranes by proteins that form the translocon pore complex (Figure 1.2) (Mota and Cornelis, 2005; Galán *et al.*, 2014). The T3SS works in conjunction with several cytoplasmic components in order to specifically

select substrates in sequential order, and to deliver them directly across the bacterial envelope into the cytosol of eukaryotic cells (Galán *et al.*, 2014).

While many of the proteins involved in the assembly of the injectisome are well conserved among the different T3SSs, the primary structure of T3S effectors usually do not show significant identity to proteins found in other organisms and also do not show significant homology among each other. Furthermore, the lack of a clear secretion signal distinguishes T3SSs from other protein secretion systems where an obvious signal is evident on the basis of the protein sequence, structure or special properties. Nevertheless, experiments done using hybrid proteins containing parts of *Yersinia enterocolitica* Yop effectors fused to reporter proteins led to the discovery that the secretion recognition domain normally lies within the first 30 amino acids of the T3S substrate (Michiels and Cornelis, 1991; Sory *et al.*, 1995). Even though, comparison of the different N-termini sequences of these effectors did not reveal sequence similarity, suggesting that the secretion signal could be conformational. Although for a while some controversy existed regarding whether the T3S signal was encoded in the peptide sequence or on the messenger RNA sequence, now it is generally believed that the secretion signal resides in the first 30 amino acids of the effector (Anderson and Schneewind, 1997; Lloyd *et al.*, 2001; Lloyd *et al.*, 2002; Page and Parsot, 2002; Galán *et al.*, 2014). In addition, bioinformatics analyses of the amino acid composition and secondary structure of the first 100 amino acids of the N-termini of known T3S substrates from *Salmonella*, *Yersinia*, *E. coli* and *Pseudomonas* (animal and plant pathogens) suggested the existence of a conserved, though highly variable, secretion signal in that region (Arnold *et al.*, 2009; Löwer and Schneider, 2009; Samudrala *et al.*, 2009; Hovis *et al.*, 2013; Wang *et al.*, 2013). Special features were also observed as enrichments of specific

amino acids (such as serine, proline and threonine) and depletion of leucine. The data accumulated allowed also the development of T3S signal prediction tools and to create lists of novel putative T3S substrates from several bacterial genomes, including *Chlamydia* organisms. However, most of these predictions remain to be tested. Despite the illusive nature of the T3S signal, one important characteristic is that the signal is universal and T3S substrates from one system can be recognized and secreted by other T3SSs (Anderson *et al.*, 1999).

1.4.3 The chlamydial T3S apparatus

The first evidence of the existence of a structure similar to the T3SS in *Chlamydia* was probably given by Matsumoto and colleagues in 1973, about 20 years before the first injectisomes from *Salmonella* were seen by electron microscopy (Matsumoto, 1973; Kubori, 1998). At that time electron micrographs of EBs and RBs from *C. psittaci* revealed structures that were called “rosettes”, which were very similar to the now known structure of the injectisome. Very recent electron microscopy studies complemented those observations and showed a significant structural similarity with the T3SS apparatus of *Yersinia*, confirming the presence of T3SS in *Chlamydia* (Dumoux *et al.*, 2014; Nans *et al.*, 2014). However, early in 1998, with the first available chlamydial genome, it had been already hypothesized that *Chlamydia* could possess a T3SS (Stephens *et al.*, 1998). Nevertheless, while in the majority of the bacteria the genes encoding injectisome and translocon proteins are clustered in chromosomal pathogenicity islands or in plasmids, in *Chlamydia*, the genes were found in four distinct conserved genomic clusters (Peters *et al.*, 2007). Focusing on chlamydial genes encoding proteins with amino acid similarity to chaperones, injectisome components, and a secreted protein (CopN) from other T3SSs, expression analyses supported the genomic analyses and established

that *Chlamydia* have an active T3SS throughout the developmental cycle (Fields *et al.*, 2003). In this work, the authors also observed that CopN could be secreted by the *Yersinia* T3S apparatus, indicating that chlamydial T3S substrates could be identified by using other bacteria as heterologous hosts. The finding of a functional T3SS was further confirmed by genome wide transcriptomic analyses during the developmental cycle and mass spectrometry studies performed in EBs and RBs (Vandahl *et al.*, 2001; Belland *et al.*, 2003; Brunelle *et al.*, 2004; Skipp *et al.*, 2005). In another study, the authors could reconstitute to some extent the *Chlamydia* T3S apparatus by generating *E. coli* expressing all genes of the chlamydial T3SS, providing a platform for further structural studies (Bao, Beatty, *et al.*, 2012). Although these studies recognized that a T3SS is active throughout the chlamydial developmental cycle, it was also important to determine the importance of T3S during *Chlamydia* infection of host cells. Experiments using molecule inhibitors known to block T3S in other bacteria were performed in *Chlamydia* to suggest that the T3SS is essential in all stages of the developmental cycle, including intracellular replication and infectivity (Muschiol *et al.*, 2006; Wolf *et al.*, 2006; Muschiol *et al.*, 2009). It was also demonstrated that chlamydial EBs can be induced to secrete T3S effectors in special inducing conditions and in the absence of host cells (Jamison and Hackstadt, 2008).

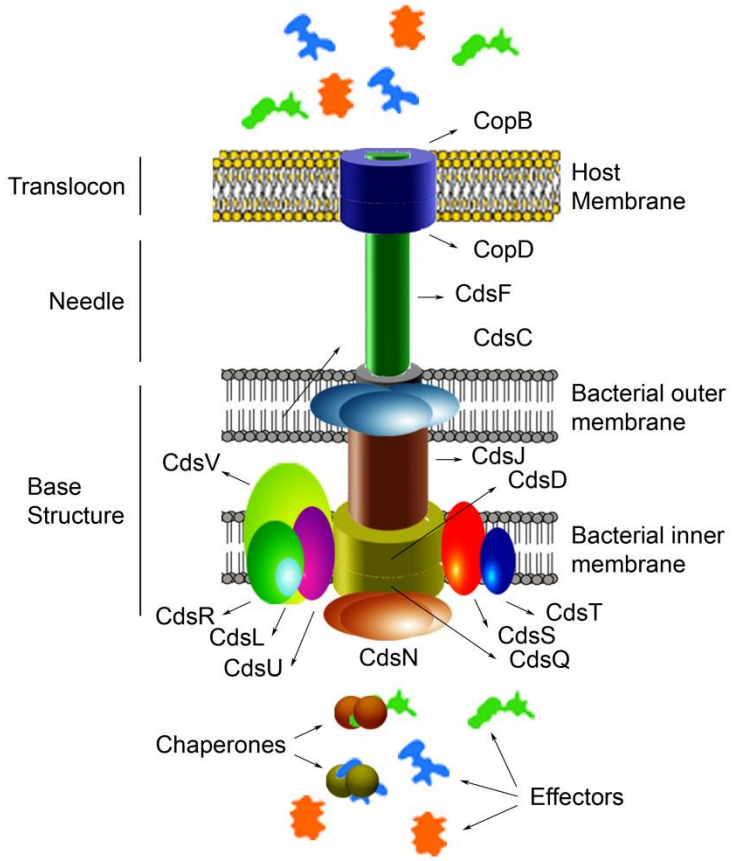


Figure 1.2 – Predicted assembly of the T3SS in *Chlamydia*. The T3S apparatus is a needle-like structure protruding from the bacterial surface that enables the translocation of virulence proteins (effectors) from the bacterial cytoplasm into the host cell cytosol or membranes. Adapted from (Peters *et al.*, 2007; Spaeth *et al.*, 2009).

1.4.4 T3S translocators of Chlamydia

Translocation of T3S effector proteins across a eukaryotic cell membrane requires a set of proteins, known as “translocators”, that are also T3S substrates. As described above, these proteins form a translocation pore, the translocon, in a host cell membrane and together with the injectisome needle enable the formation of a continuous channel between the bacteria and the target cell (Figure 1.2). Generally, there are three translocators: two of them have clear hydrophobic

domains and insert into the target cell membrane, and the other is a hydrophilic protein that forms a complex at the needle tip, possibly functioning as an assembly platform for the formation of the pore by the hydrophobic translocators (Mueller *et al.*, 2008).

In *Chlamydia*, there are two copies of each of the hydrophobic translocator proteins (CopB1 and CopB2, and CopD1 and CopD2), which might function in the formation of specific translocons in the host cell membrane and in the inclusion membrane (Stephens *et al.*, 1998; Peters *et al.*, 2007). Until now, it is unclear which *Chlamydia* protein might form the needle tip complex and function as the hydrophobic translocator. A biophysical characterization of CT584 from *C. trachomatis* suggested it could be the needle complex protein (Markham *et al.*, 2009), but subsequent studies indicated that this might not be the case (Spaeth *et al.*, 2009; Stone *et al.*, 2012; Barta *et al.*, 2013; Pais *et al.*, 2013).

1.4.5 T3S chaperones of Chlamydia

The presence of specific cytoplasmic chaperones is also a hallmark of T3SSs. They can been divided into three classes: the ones that bind effectors (class I; the best studied), the pore forming translocators (class II), or the subunits of the injectisome (class III) (Page and Parsot, 2002; Feldman and Cornelis, 2003). T3S chaperones share low sequence similarity between each other but display a conserved three-dimensional structure. They bind to one or several partners, thus preventing premature, or incorrect, intra- or intermolecular interactions that would occur in the absence of the chaperone. T3S chaperones participate in the assembly and regulation of the injectisome and assist in the secretion of effectors, possibly establishing a hierarchy of secretion (Figure 1.2). They are characterized by a low molecular mass

(approximately 15 kDa), have acidic pI and normally form dimers. The chaperone binding domain within effectors is generally localized downstream from the secretion signal (Page and Parsot, 2002; Feldman and Cornelis, 2003).

There are already several proteins acting as T3S chaperones described in *Chlamydia*: Mcsc, Slc1, Scc1, Scc2, Scc3, Scc4, and CT584. Mcsc binds to and stabilizes at least Cap1 and the Inc protein CT618 (Spaeth *et al.*, 2009). Slc1 was described to enhance the secretion of the T3S effectors Tarp, CT694, TepP and the T3S substrate CT695 (Brinkworth *et al.*, 2011; Pais *et al.*, 2013; Chen *et al.*, 2014). Scc1 and Scc4 are the chaperones of CopN (Silva-Herzog *et al.*, 2011). CT584, initially described as the needle tip protein (Stone *et al.*, 2012), as discussed above, was shown to bind, stabilize and promote secretion of the putative T3S substrate CT082 (Pais *et al.*, 2013), suggesting it could function as a T3S chaperone. It is possible that CT584 might have additional functions, which is not unusual for T3S chaperones. Finally, Scc2 and Scc3 are likely class II chaperones possibly promoting secretion of Chlamydia translocator proteins (CopB1 and CopB2) (Fields *et al.*, 2005).

1.4.6 T3S effectors of Chlamydia

It is predicted that approximately 5–10% of the *Chlamydia* genomes encode for T3S effectors (Arnold *et al.*, 2009; Samudrala *et al.*, 2009). Before the discovery of genetic methodologies to manipulate *Chlamydia*, chlamydial effector proteins were mainly searched for by:

- Constructing *Saccharomyces cerevisiae* strains expressing *Chlamydia*-specific genes in order to identify chlamydial proteins that impair various yeast cellular functions or that display tropism towards eukaryotic organelles (Sisko *et al.*, 2006);

- Exploiting the universality of the T3S signal and using *Salmonella*, *Shigella*, or *Yersinia* as heterologous host bacteria carrying T3SSs (Fields and Hackstadt, 2000; Clifton *et al.*, 2004; Ho and Starnbach, 2005; Subtil *et al.*, 2005; Chellas-Géry *et al.*, 2007; Hower *et al.*, 2009; Pennini *et al.*, 2010; Dehoux *et al.*, 2011; Muschiol *et al.*, 2011; Furtado *et al.*, 2013; Hovis *et al.*, 2013; Pais *et al.*, 2013);
- Using T3S signal prediction tools to identify novel putative T3S substrates (Arnold *et al.*, 2009; Löwer and Schneider, 2009; Samudrala *et al.*, 2009; Hovis *et al.*, 2013; Wang *et al.*, 2013);
- Using specific antibodies enabling the detection of translocation into host cells of chlamydial proteins predicted to be effectors (Li, Chaoqun Chen, *et al.*, 2008; Gong *et al.*, 2011; Lei *et al.*, 2011);
- Using T3S molecule inhibitors to confirm secretion of an effector by the T3SS (Hobolt-Pedersen *et al.*, 2009; Gong *et al.*, 2011);
- Using mass spectrometry analyses to identify the battery of proteins secreted into the host cytosol during infection (Kleba and Stephens, 2008);
- Using *in vitro* biochemical assays directed to test specific properties of predicted chlamydial effectors (Misaghi *et al.*, 2006; Le Negrate *et al.*, 2008).

The chlamydial T3S effectors discovered to date, specifically the ones that are known to be translocated across the inclusion membrane in infected cells and/or are known to interfere with host cell functions, are summarized in Table 1.1.

Table 1.1 – List of chlamydial T3S effectors known to be translocated across the inclusion membrane and/or known to interfere with host cell functions.

T3S Effector[a]	Subcellular localization	Described target(s)	Feature or proposed function	Species specificity[b]	Reference
CT089/CopN	Inclusion membrane	?	?	*Chlamydia*	(Fields and Hackstadt, 2000)
CT456/Tarp	Cytosol	Actin	Polymerizes actin at the site of entry	*Chlamydia*	(Clifton *et al.*, 2004; Jewett *et al.*, 2006; Lane *et al.*, 2008)
CT529/Cap1	Inclusion membrane	?	Stimulates a CD8+ T cell response	Chlamydia	(Fling et al., 2001)
CT620	Nuclei	?	?	*Chlamydia*	(Muschiol *et al.*, 2011)
CT621	Nuclei	?	?	*Chlamydia*	(Hobolt-Pedersen *et al.*, 2009; Muschiol *et al.*, 2011)
CT622	Cytosol	?	?	*Chlamydia*	(Gong *et al.*, 2011)
CT694	Cytosol	Ahnak	Induces actin-stress fibers	*C. trachomatis* *C. muridarum*	(Hower *et al.*, 2009; Bullock *et al.*, 2012)
CT695	Cytosol	?	?	*Chlamydia*	(Mueller and Fields, 2015)
CT711	Nuclei	?	?	*Chlamydia*	(Muschiol *et al.*, 2011)
CT737/NUE	Nuclei	Histones	Acts as histone methyltransferase	*Chlamydia*	(Pennini *et al.*, 2010)
CT847	?	GCIP	Interferes with cell cycle checkpoints	*Chlamydia*	(Chellas-Géry *et al.*, 2007)
CT875/TepP	Cytosol	?	Acts in signaling pathways of regulation of innate immune responses	*C. trachomatis*	(Chen *et al.*, 2014)
CPn0483/*Chla*OTU	?	Ubiquitin NDP52	Acts in the clearance of ubiquitin at the invasion sites	*Chlamydia* except *C. trachomatis* and *C. muridarum*	(Furtado *et al.*, 2013)
G5Q-0070/SINC	Inner nuclear membrane	ELYS Lamin B1 Emerin	Alters the nuclear envelope function	*C. psittaci*	(Mojica *et al.*, 2015)
Inc CT101	Inclusion membrane	?	Co-localizes with proteins of the Src family	*Chlamydia* except *C. muridarum*	(Mital *et al.*, 2010)

T3S Effector[a]	Subcellular localization	Described target(s)	Feature or proposed function	Species specificity[b]	Reference
Inc CT115/IncD	Inclusion membrane	CERT VapB	Recruits ER proteins to the inclusion	*Chlamydia* except *C. pneumoniae*	(Derré *et al.*, 2011; Agaisse and Derré, 2014)
Inc CT116/IncE	Inclusion membrane	SNX5/6	Disrupts retromer function	*C. trachomatis* *C. muridarum*	(Mirrashidi *et al.*, 2015)
Inc CT118/IncG	Inclusion membrane	14-3-3-β	?	*C. trachomatis* *C. muridarum*	(Scidmore and Hackstadt, 2001)
Inc CT119/IncA	Inclusion membrane	Vamp3 Vamp8	Promotes fusion of membranes	*Chlamydia* except *C. caviae* *C. pneumoniae*	(Hackstadt *et al.*, 1999; Delevoye *et al.*, 2008)
Inc CT222	Inclusion membrane	?	Co-localizes with proteins of the Src family	*C. trachomatis*	(Mital *et al.*, 2010)
Inc CT223/IPAM	Inclusion membrane	CEP170	Assembles microtubule superstructures crucial to maintain the inclusion shape	*C. trachomatis* *C. muridarum* *C. pneumoniae*	(Dumoux *et al.*, 2015)
Inc CT228	Inclusion membrane	MYPTI	Regulates the extrusion mechanism to exit from the host cell	*C. trachomatis* *C. muridarum* *C. caviae*	(Mital *et al.*, 2015)
Inc CT229	Inclusion membrane	Rab4	?	*C. trachomatis* *C. muridarum*	(Rzomp *et al.*, 2006)
Inc CT232/IncB	Inclusion membrane	?	Co-localizes with proteins of the Src family	*Chlamydia*	(Mital *et al.*, 2010)
Inc CT813/InaC	Inclusion membrane	14-3-3-β ARF1	Assembles microtubule superstructures crucial to maintain the inclusion shape	*C. trachomatis* *C. muridarum*	(Chen *et al.*, 2006; Kokes *et al.*, 2015)
Inc CT850	Inclusion membrane	DYNLT1	Promotes the correct positioning of the inclusion at the MTOC	*Chlamydia*	(Mital *et al.*, 2015)
Inc Cpn0517	Inclusion membrane	Act1	Inhibits NF-κB activation	*C. pneumoniae*	(Wolf *et al.*, 2009)
Inc Cpn0127	?	Caprin2	Inhibits apoptosis	*C. pneumoniae*	(Flores and Zhong, 2015)
Inc G5Q-0512/IncB	Inclusion membrane	Snapin Dync1i1	Connects the inclusions with the microtubule network	*C. psittaci*	(Böcker *et al.*, 2014)

[a]Based on the protein sequence annotation of *C. trachomatis* strain D/UW3, *C. pneumonia* CWL029, and *C. psittaci* CAL10. The Inclusion membrane (Inc) proteins are indicated as Inc followed by their genome sequence annotation.

[b]The *Chlamydia* species where a homolog of the protein is found, among *C. trachomatis, C. muridarum*, *C. caviae*, *C. felis*, and *C. pneumoniae* (Lutter *et al.*, 2012; Mirrashidi *et al.*, 2015).

1.4.6.1 Chlamydia *inclusion membrane (Inc) proteins*

The Inc proteins are a particular and important class of *Chlamydia* T3SS effectors. It is known since 1995 that *Chlamydia* secrete proteins that localize at the inclusion membrane (Rockey *et al.*, 2002). Among these proteins, several have a characteristic bi-lobed hydrophobic domain (possibly enabling them to intercalate in the inclusion membrane) and are collectively known as Inc proteins (Figure 1.3) (Bannantine *et al.*, 2000).

Inc proteins have been identified by searching the genome for the unique signature of the bi-lobed hydrophobic domain (Bannantine *et al.*, 2000). Depending on the species, *Chlamydia* encode a vast number of Incs: ~60 in *C. trachomatis*, ~60 in *C. muridarum*, ~85 in *C. caviae*, ~80 in *C. felis*, and ~100 in *C. pneumoniae* (Bannantine *et al.*, 2000; Dehoux *et al.*, 2011; Lutter *et al.*, 2012). The Inc proteins appear to be specific of *Chlamydiae*, as they were also identified in the endosymbiont of free-living amoebae *Protochlamydia amoebophila* but not in other bacteria (Griffiths *et al.*, 2006; Heinz *et al.*, 2010). The entire collection of Inc proteins is not conserved among all species, which indicate that the diversity among these proteins likely contributes to the subtle, species-specific differences observed in host range. About 25 Inc proteins are conserved among all species (Lutter *et al.*, 2012).

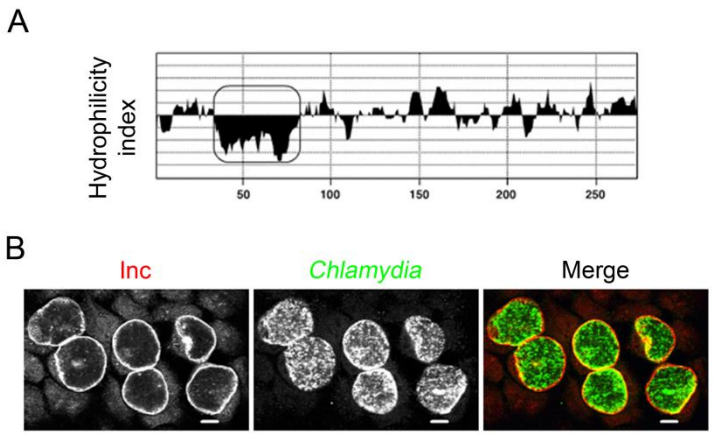


Figure 1.3 – The inclusion membrane (Inc) proteins possess a bi-lobed hydrophobic motif that targets them to the inclusion membrane. (A) Typical hydrophilicity profile of an Inc protein. Reprinted from Cellular Microbiology; Vol. 2; John P. Bannantine, R. S. Griffiths, W. Viratyosin, Wendy J. Brown, Daniel D. Rockey; A secondary structure motif predictive of protein localization to the chlamydial inclusion membrane; 35-47; 2000; with permission from John Wiley and Sons. (B) HeLa 229 mammalian cells infected with *C. trachomatis* L2/434 and visualized by fluorescence microscopy. Cells were immunolabeled with anti-MOMP (green) for *Chlamydia* and anti-Inc (red) for the Inc. Scale bar is 10 µm.

Given that Inc proteins have the distinctive localization at the inclusion membrane, most probably facing the host cytosol, they are believed to be key mediators of *Chlamydia*-host interactions. Nevertheless, not all predicted Inc proteins were yet validated as the distinctive localization at the inclusion membrane was only experimentally confirmed for ~25 of them (Cortes *et al.*, 2007; Li, Chaoqun Chen, *et al.*, 2008; Mital *et al.*, 2010; Dehoux *et al.*, 2011; Flores and Zhong, 2015). A recent study used the newly developed genetic tools to identify 10 additional Inc proteins localizing at the inclusion membrane (Weber *et al.*, 2015).

Inc proteins lack predicted signal sequences required for secretion by other transport mechanisms such as the general secretory pathway. This led to the hypothesis that they could be translocated to the host cell by a T3SS. Heterologous bacteria harboring T3SSs, such as *Shigella*

flexneri have been used to corroborate this hypothesis (Subtil *et al.*, 2001; Subtil *et al.*, 2005; Dehoux *et al.*, 2011).

The molecular function of many Inc proteins is still unknown; nevertheless, considerable advances have been made for several of them (Table 1.1). Inc CT115/IncD was described to be enriched in MCSs and interact with the host lipid transfer protein CERT and VAPB. As mentioned above in section 1.3.3.5, this interaction enables the transfer of ER content into the inclusion (Derré *et al.*, 2011; Agaisse and Derré, 2014; Agaisse and Derré, 2015).

Inc CT118/IncG is known to be phosphorylated during infection and was described to bind to the phosphoserine-binding protein 14-3-3-β, involved in signal-transduction pathways (Scidmore and Hackstadt, 2001).

Inc CT119/IncA mediates the homotypic fusion of the inclusions during infection by *Chlamydia* through the action of a SNARE domain, and possibly by the interaction with host VAMP3 and VAMP8 (Hackstadt *et al.*, 1999; Delevoye *et al.*, 2008; Ronzone and Paumet, 2013; Ronzone *et al.*, 2014). In a recent study, where the authors documented the first inactivation of a chromosomal gene in *Chlamydia*, it was demonstrated that the lack of IncA production renders non-fusogenic inclusions in cells infected with this mutated strain (Johnson and Fisher, 2013).

There is a group of Inc proteins that localize in microdomains at the inclusion membrane (Incs CT101, CT222, CT232/IncB, and CT850). These microdomains are localized at the point of contact of the centrosome with the inclusion and are also enriched in a family of host Src kinases that interact with many host proteins, modifying them by phosphorylation of tyrosine residues. It is suggested that these microdomains may function as a platform for interactions between the chlamydial inclusion and the host (Mital *et al.*, 2010).

Inc CT850 was recently described to interact with the dynein light chain (DYNLT1) and to contribute to the correct migration of the inclusion along microtubules to the centrosome region (Mital *et al.*, 2015).

As mentioned above in section 1.3.3.7, the Inc proteins CT223/IPAM and CT813/InaC are capable of reprogramming the host microtubule network. IPAM interacts with the host centrosomal protein CEP170 (Dumoux *et al.*, 2015), and InaC binds host ARF1 and 14-3-3 proteins (Kumar and Valdivia, 2008; Kokes *et al.*, 2015).

Inc CT228 was shown to interact with the myosin phosphatase target subunit 1 (MYPT1), an element of the myosin phosphatase pathway, and as mentioned above in section 1.3.3.9, this interaction favors the extrusion mechanism (Lutter *et al.*, 2013).

The *C. pneumoniae*-specific Inc Cpn0127 was described to interact with the host cell cytoplasmic activation/proliferation-associated protein 2 (Caprin2), indicating an active role in the manipulation of host Wnt signaling pathway for enhancing the chlamydial anti-apoptotic activity (Flores and Zhong, 2015).

Although Inc CT232/IncB is present in all *Chlamydia* species, the homolog of this protein of *C. psittaci*, Inc G5Q-0512/IncB, has only ~50% amino acid identity to CT232/IncB of *C. trachomatis*. Inc G5Q-0512/IncB was shown to interact with host Snapin and dynein, indicating that Snapin connects the inclusion with the microtubule network by interacting with both IncB and dynein (Böcker *et al.*, 2014). Interestingly, this interaction was not observed with CT232/IncB of *C. trachomatis*, highlighting the species specificity of the function of this protein (Böcker *et al.*, 2014).

Recently, the development of a systematic affinity purification-mass spectrometry (AP-MS) approach led to the identification of a

comprehensive Inc-host interactome. In this work, the authors could find high-confidence host targets for 38 out of 58 *C. trachomatis* Inc proteins (Mirrashidi *et al.*, 2015). Of the 20 validated Incs (experimentally confirmed to localize at the inclusion membrane) used in this study, the authors identified host targets for 15, of which two had been already described by other methods (Incs CT115/IncD and CT228) (Derré *et al.*, 2011; Lutter *et al.*, 2013; Mirrashidi *et al.*, 2015). In this study, Inc proteins were found to target several host structures, organelles and pathways, including the ER, mitochondria, vesicle transport, cell cycle/division, apoptosis, and protein transport. The biological significance of the interactions found was further investigated for one of the Inc proteins. Inc CT116/IncE was described to contribute to the disruption of retromer function, by interacting with SNX5 and SNX6, revealing that retromer restricts *Chlamydia* infection (Mirrashidi *et al.*, 2015). Nevertheless, the other identified host interacting partners still lack experimental validation.

Two different studies highlighted an alternative role for Inc proteins. In one study Incs were described to interact with each other and in another study Incs were proposed to form complex membrane structures (Mital *et al.*, 2013; Gauliard *et al.*, 2015). This led to the hypothesis that instead of having a direct host interacting partner, some Inc proteins could form specific multi-protein complexes in the inclusion membrane to act as scaffolds for host interactions or to confer structural stability to the inclusion (Mital *et al.*, 2013; Gauliard *et al.*, 2015).

1.4.6.2 Other Chlamydia *T3S effectors*

The Inc proteins are not the only chlamydial proteins directed to the membrane of the inclusion. *Chlamydia* secrete two T3S effectors (CT089/CopN and CT529/Cap1) that were also described to have that

distinct localization, and do not possess the typical architecture of an Inc protein (Table 1.1) (Fields and Hackstadt, 2000; Fling *et al.*, 2001). However, nothing is known about the molecular function of these two proteins.

Among the remaining T3SS effectors known so far (Table 1.1), Tarp is crucial for the chlamydial entry step. As soon as EBs attach to the plasma membrane, Tarp is secreted to the host cytosol and it is immediately phosphorylated. Then, actin is recruited to the site of attachment in a WAVE2 and Arp2/3-dependent manner, where it is polymerized by Tarp to promote bacterial endocytosis (Clifton *et al.*, 2004; Jewett *et al.*, 2006; Lane *et al.*, 2008). The TepP is also secreted early in infection to induce signaling cascades that modulate the immune response against the bacteria (Chen *et al.*, 2014).

CT694 is another T3SS effector secreted in the invasion process and is known to interact with the human protein Ahnak, affecting the formation of host-cell actin stress fibers (Hower *et al.*, 2009; Bullock *et al.*, 2012). The homolog of this protein of *C. psittaci*, G5Q-0070/SINC, has a low amino acid sequence identity to CT694 of *C. trachomatis* (Mojica *et al.*, 2015). SINC (secreted inner nuclear membrane–associated *Chlamydia* protein) was reported to target the nuclear envelope of *C. psittaci*–infected cells and to interact with proteins of the nuclear membrane, including the nucleoporin ELYS, lamin B1, emerin, MAN1, LAP1, and LBR. Therefore, in *C. psittaci*, SINC is believed to control other processes such as nuclear structure and chromatin organization (Mojica *et al.*, 2015).

Chlamydia has 4 more proteins that are targeted to the host cell nucleus: CT620, CT621, CT711, and CT737/NUE (Hobolt-Pedersen *et al.*, 2009; Pennini *et al.*, 2010; Muschiol *et al.*, 2011). CT737 is named nuclear effector (NUE) and contains a domain of histone methyltransferase.

Microscopy experiments revealed that during infection NUE is translocated to the host cell nucleus and *in vitro* enzymatic assays confirmed the activity of NUE as a histone methyltransferase (Pennini *et al.*, 2010). For the other three proteins (CT602, CT621, and CT711), there is evidence of a T3S signal encoded in their N-termini and specific antibodies raised against these proteins showed that they are present in the nuclei of *Chlamydia*-infected cells (Hobolt-Pedersen *et al.*, 2009; Muschiol *et al.*, 2011).

CT847 was found to interact with the mammalian Grap2 cyclin D-interacting protein (GCIP). During infection, reduced levels of GCIP were observed, indicating that this protein might be targeted for degradation. Though, the authors could not directly implicate the interaction of the two proteins on the observed *Chlamydia*-mediated effect of reduced levels of GCIP (Chellas-Géry *et al.*, 2007).

1.4.7 Other Chlamydia *effectors*

Although the majority of the effectors known so far are translocated to the host cell by the T3SS, *Chlamydia* also secrete proteins by other transport systems. These include the CT166/cytotoxin, CT311, CT441/Tsp (tail specific protease), CT795, CT798/GlgA (glycogen synthase), CT823/HtrA (high temperature requirement protein A), CT858/CPAF (chlamydial protease-like activity factor), and the plasmid ORF Pgp3 (Lad, Yang, *et al.*, 2007; Li, Ding Chen, *et al.*, 2008; Thalmann *et al.*, 2010; Lei *et al.*, 2011; Qi *et al.*, 2011; Wu *et al.*, 2011; Zhong, 2011; Lei *et al.*, 2013; Lu *et al.*, 2013; Bothe *et al.*, 2015). Furthermore, other proteins highlighted in previous studies could be T3S substrates but this has never been confirmed, and therefore, we will describe their known functions here (Kumar *et al.*, 2006; Misaghi *et al.*, 2006; Le Negrate *et al.*, 2008; Taylor *et al.*, 2010).

The most extensively studied non-T3S substrates are the CT166/cytotoxin and the CT858/CPAF. The CT166/cytotoxin is present in EBs, prior to infection, and is detected in the host-cell cytosol in the first stages of infection. The cytotoxin was reported to cause actin re-organization events when expressed in uninfected mammalian cells (Carlson *et al.*, 2004; Thalmann *et al.*, 2010; Bothe *et al.*, 2015). As the same effect was observed with the *Clostridium difficile* Toxin B (TcdB), the authors hypothesized that the chlamydial cytotoxin could also act in actin re-organization, possibly through the interaction with Rac1 (Carlson *et al.*, 2004; Thalmann *et al.*, 2010; Bothe *et al.*, 2015).

CPAF is an extensively characterized protease with numerous potential host substrates of importance to *Chlamydia* infection (Zhong, 2011). However, there has been some controversy around this protein and whether the identified substrates indeed constitute bona fide targets of proteolysis in infected cells or correspond to artifacts of post-lysis degradation (Chen *et al.*, 2012). In a recent study, where CPAF mutants were generated, the authors reported that some of the effects previously associated with CPAF were indeed independent of this protein (Snavely *et al.*, 2014). Nevertheless, they could confirm other targets as being CPAF-dependent and suggested a role for CPAF late in infection, possibly during the stages leading to the disruption of the host cell prior to the release of EBs (Snavely *et al.*, 2014).

C. trachomatis secretes at least three proteins into host cells that localize at distinct lipid droplet-like structures surrounding the inclusion (Kumar *et al.*, 2006). These proteins were named lipid droplet associated proteins (CT156/Lda1, CT163/Lda2 and CT473/Lda3), and encode phospholipase domains, present in lipid-modifying enzymes. They are suggested to participate in the recruitment of host lipids into the inclusion (Beatty, 2006; Kumar *et al.*, 2006; Beatty, 2008; Cocchiaro *et al.*, 2008; Saka *et al.*, 2015). Since they lack a predicted signal for

known protein transport systems, it is believed that they are secreted across the inclusion membrane through the T3SS (Kumar *et al.*, 2006).

Bioinformatics analyses of the *C. trachomatis* genome revealed other hypothetical proteins that could function as T3S effectors, as the case of CT153/MACPF (Samudrala *et al.*, 2009; Taylor *et al.*, 2010). This protein contains a MACPF domain that is used by different parasites to facilitate the egress from cells (Kadota *et al.*, 2004; Kafsack *et al.*, 2009). This suggests that *Chlamydia* could also employ a similar process of producing pores in membranes to facilitate disruption of the inclusion and the host cell (Rosado *et al.*, 2008; Taylor *et al.*, 2010). Another possible role for this protein could be in the import of host lipids into the inclusion, in conjunction with the Lda proteins, as their genes are located adjacently in the genome of *C. trachomatis* (Rosado *et al.*, 2008; Taylor *et al.*, 2010).

1.5 Other intracellular pathogens

There are pathogenic bacteria that successfully thrive in the host cytosol, as the case of *S. flexneri* and *Listeria monocytogenes* (Cossart and Toledo-Arana, 2008; Schroeder and Hilbi, 2008). In general, these bacteria avoid killing by the host through inactivation of the autophagic pathway (Kumar and Valdivia, 2009; Asrat *et al.*, 2014). Nevertheless, the majority of the intracellular pathogens reside within membrane-bound compartments. Those are typically the case of *Chlamydia*, *Mycobacterium tuberculosis*, *Legionella pneumophila*, *Brucella abortus*, *Coxiella burnetii* and *Salmonella enterica* (Fields and Hackstadt, 2002; Vergne *et al.*, 2004; van Schaik *et al.*, 2013; Xu and Luo, 2013; Celli, 2015; LaRock *et al.*, 2015). This intracellular lifestyle might be advantageous for the bacteria, as this environment enables them to escape more easily from the host immune response soon after invasion (Kumar and Valdivia, 2009; Hilbi and Haas, 2012; Asrat *et al.*, 2014).

1.5.1 Vacuole interactions with the host by intravacuolar pathogens

Although remaining within a vacuole inside the host cell for the duration of the intracellular infectious cycle, each intravacuolar pathogen evolved different approaches to modify its compartment, and to manipulate the host endocytic and the secretory pathway to prevent killing by the host (Figure 1.4) (Kumar and Valdivia, 2009).

1.5.1.1 Salmonella enterica

S. enterica replicates in epithelial cells inside the *Salmonella* containing vacuole (SCV) (LaRock *et al.*, 2015). Maturation of the SCV involves transient interactions with early endosomes, with recruitment of early

endosomal markers that are subsequently replaced with late endosomal markers and some lysosomal markers. Despite the direct fusion of the SCV with lysosomes, the endocytic pathway is interrupted and bacteria are able to endure throughout the rest of the infectious cycle (LaRock *et al.*, 2015).

1.5.1.2 Mycobacterium tuberculosis

Mycobacterium tuberculosis is highly adapted to infect macrophages, and manages to escape killing by the host through the establishment of the *Mycobacteria* pathogen vacuole (MPV) (Vergne *et al.*, 2004). The MPV deviates from the canonical endocytic pathway between early and late endosomes and is characterized by an incomplete acidification of its lumen, and the absence of lysosomal hydrolases and Rab proteins (Vergne *et al.*, 2004).

1.5.1.3 Coxiella burnetii

C. burnetii has a tropism for professional phagocytes and replicates within the *Coxiella*-containing vacuole (CCV) (van Schaik *et al.*, 2013). Unlike other bacterial pathogens, where maturation of the vacuole involves the arrest in early or late endosomes, or lysosomes, the CCV follows the endocytic pathway till the end. The main difference between the generation of the CCV and lysosomes is the time at which lysosomal markers are delivered in each compartment. A delay in the transport of these markers to the CCV allows the expansion of the vacuole and the accumulation of further cellular material that benefits bacterial survival (van Schaik *et al.*, 2013).

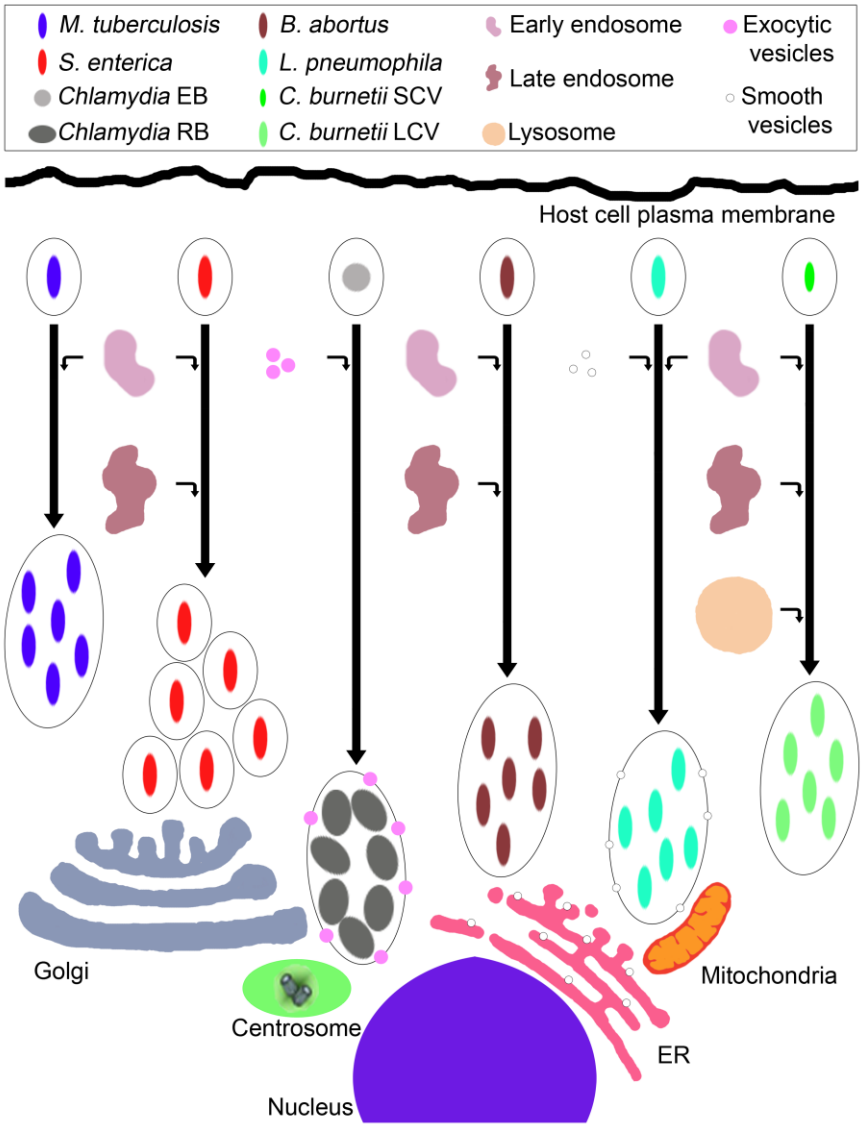


Figure 1.4 - Vacuole interactions with the host by intravacuolar pathogens. The *M. tuberculosis* vacuole interrupts the endocytic pathway between early and late endosomes. *S. enterica* avoids the acquisition of lysosomal markers and establishes its vacuole near the Golgi. The *Chlamydia* inclusion does not interact with the endocytic pathway, and instead fuses with exocytic vesicles and migrates towards the centrosomes. The *B. abortus* vacuole acquires markers of early and late endosomes but then change into an endoplasmic reticulum (ER)-derived vacuole. The *L. pneumophila* vacuole interacts with smooth vesicles from the ER, originating an ER-derived vacuole that recruits other organelles such as mitochondria. The *C. burnetii* vacuole follows the endocytic pathway till the end, acquiring all the characteristics of lysosomes. EB – elementary body; RB – reticulate body; SCV – *Coxiella* small cell variant; LCV – *Coxiella* large cell variant. Adapted from (Hilbi and Haas, 2012; Asrat *et al.*, 2014).

1.5.1.4 Brucella abortus

B. abortus infects mainly professional phagocytes, and upon entry of a host cell, bacteria are engulfed in the *Brucella*-containing vacuole (BCV) (Celli, 2015). The BCV acquires markers of early and late endosomes but not of lysosomes. Instead, the BCV intercepts the early secretory pathway at the ER and is drastically modified from a phagosome to an ER-derived organelle, by following the autophagic pathway (Celli, 2015).

1.5.1.5 Legionella pneumophila

L. pneumophila survives and replicates within macrophages in the *Legionella*-containing vacuole (LCV) (Xu and Luo, 2013). Upon invasion of a host cell, the LCV rapidly deviates from the endocytic pathway to the secretory pathway. Instead of acquiring markers of late endosomes, the LCV is covered with smooth vesicles originating from the ER. During the remaining of the infectious cycle, the LCV becomes more enriched in ER material and other organelles such as mitochondria, conferring not only protection but also providing the bacteria with nutrients (Xu and Luo, 2013).

1.5.2 Interference with host cell pathways by intravacuolar pathogens

Alongside with the different vacuolar biogenesis, each intravacuolar pathogenic bacterium employs different virulence mechanisms and effectors to accomplish the above mentioned modifications (Kumar and Valdivia, 2009).

1.5.2.1 Salmonella enterica

S. enterica sequentially employs two distinct T3SSs (LaRock *et al.*, 2015). The *Salmonella* pathogenicity island-1-encoded T3SS (SPI-1 T3SS) delivers effectors necessary for the invasion process, to induce bacterial internalization and to promote the first steps of the SCV maturation. The SPI-2 T3SS delivers effectors associated with bacterial replication and the correct positioning and maintenance of the SCV at the perinuclear region. Although each T3SS has its own set of effectors, there are a few that are shared by both T3SSs (LaRock *et al.*, 2015).

1.5.2.2 Mycobacterium tuberculosis

M. tuberculosis secretes proteins through a T7SS (ESX-1 to -5) (Simeone *et al.*, 2009). Despite being essential for virulence, the ESX systems are also suggested to be involved in more general aspects of the physiology of the bacteria. The ESX-1 is the main system related to pathogenicity and the effectors of this system are responsible for modulating apoptosis and assisting in the late stages of the MPV maturation (Simeone *et al.*, 2009).

1.5.2.3 Coxiella burnetii

C. burnetii cycles between two bacterial forms: the metabolically inactive small cell variant and the metabolically active large cell variant (van Schaik *et al.*, 2013). Transition from the small to the large cell variant coincides with the vacuole maturation. *C. burnetii* is known to encode several transport systems, still, the T4SS is the best studied (van Schaik *et al.*, 2013). Translocation of T4S effectors only occurs 8 hours after infection, indicating that maturation of the CCV is not dependent of the T4SS. To date more than 100 T4S substrates were identified and most of them encode one or more eukaryotic-like domains, suggesting that

the *C. burnetii* T4SS effectors interact with host proteins to maintain the CCV integrity and to modulate other processes such as apoptosis and proteasome-mediated degradation (van Schaik *et al.*, 2013).

1.5.2.4 Brucella abortus

B. abortus also uses a T4SS described to be essential for the BCV biogenesis and replication (Celli, 2015). The *B. abortus* T4S effectors are known to modulate host functions involved in the early secretory pathway and vesicle trafficking to the ER, enabling the bacteria to associate with this organelle and proliferate within the mature BCV (Celli, 2015).

1.5.2.5 Legionella pneumophila

L. pneumophila entirely depends on its T4SS to overcome the killing mechanisms of the host (Xu and Luo, 2013). From the moment of invasion till the egress of the bacteria, *L. pneumophila* actively secretes almost 300 T4S effectors that fulfill the needs of the bacteria to intercept the endocytic and the secretory pathways, to interfere with host death mechanisms, and to acquire nutrients (Xu and Luo, 2013). Nevertheless, this high abundance of effectors likely leads to an increased probability for functional redundancy among the effectors, which prevents further knowledge of the mechanisms by which *L. pneumophila* creates its unique compartment.

1.6 General aims and overview

This PhD thesis aimed at increasing our understanding of *Chlamydia*-host cell interactions. It focused in analyses of *C. trachomatis* Inc proteins, which were hypothesized to play a major role in host cell tropism and types of infection, and to modulate host cell processes during infection. Three main topics were addressed:

1) The identification of all Incs that are T3S substrates;
2) The contribution of Inc proteins to the tissue tropism and type of infection associated with *C. trachomatis* strains;
3) The study of the molecular function of Incs, including the search for host cell protein targets.

In Chapter 2, we used *Yersinia enterocolitica* as heterologous bacteria to identify all Incs that are T3S substrates, in an attempt to define a chlamydial protein as an Inc based on the presence of the typical hydrophobic motif and of a T3S signal. As we were unsuccessful in finding a T3S signal for three known Incs, we considered for further studies the group of all Incs. In Chapter 3 we determined the differences in the amino acid sequence of Inc proteins and the temporal expression levels of *inc genes* among different *C. trachomatis* strains, and found a subgroup of Incs possibly involved in the distinct tropism and invasiveness associated with *C. trachomatis* LGV strains. In Chapter 4 we performed a search for interacting host cell proteins for the subgroup of Inc proteins found in Chapter 3, and further characterized the interaction between the Inc CT288 and the host cell centrosomal protein CCDC146.

1.7 References

AbdelRahman, Y.M., and Belland, R.J. (2005) The chlamydial developmental cycle. *FEMS Microbiol Rev* **29**: 949–959.

Agaisse, H., and Derré, I. (2013) A *C. trachomatis* Cloning Vector and the Generation of *C. trachomatis* Strains Expressing Fluorescent Proteins under the Control of a *C. trachomatis* Promoter. *PLoS One* **8**.

Agaisse, H., and Derré, I. (2014) Expression of the effector protein IncD in *Chlamydia trachomatis* mediates recruitment of the lipid transfer protein CERT and the endoplasmic reticulum-resident protein VAPB to the inclusion membrane. *Infect Immun* **82**: 2037–2047.

Agaisse, H., and Derré, I. (2015) STIM1 Is a Novel Component of ER-*Chlamydia trachomatis* Inclusion Membrane Contact Sites. *PLoS One* **10**: e0125671.

Albrecht, M., Sharma, C.M., Dittrich, M.T., Müller, T., Reinhardt, R., Vogel, J., and Rudel, T. (2011) The transcriptional landscape of *Chlamydia pneumoniae*. *Genome Biol* **12**: R98.

Alzhanov, D.T., Weeks, S.K., Burnett, J.R., and Rockey, D.D. (2009) Cytokinesis is blocked in mammalian cells transfected with *Chlamydia trachomatis* gene *CT223*. *BMC Microbiol* **9**: 2.

Arnold, R., Brandmaier, S., Kleine, F., Tischler, P., Heinz, E., Behrens, S., *et al.* (2009) Sequence-Based Prediction of Type III Secreted Proteins. *PLoS Pathog* **5**.

Asrat, S., Jesús, D.A. de, Hempstead, A.D., Ramabhadran, V., and Isberg, R.R. (2014) Bacterial Pathogen Manipulation of Host Membrane Trafficking. *Annu Rev Cell Dev Biol* **30**: 79–109.

Bachmann, N.L., Fraser, T. a, Bertelli, C., Jelocnik, M., Gillett, A., Funnell, O., *et al.* (2014) Comparative genomics of koala, cattle and sheep strains of *Chlamydia pecorum*. *BMC Genomics* **15**: 667.

Bannantine, J.P., Griffiths, R.S., Viratyosin, W., Brown, W.J., and Rockey, D.D. (2000) A secondary structure motif predictive of protein localization to the chlamydial inclusion membrane. *Cell Microbiol* **2**: 35–47.

Bao, X., Nickels, B.E., and Fan, H. (2012) *Chlamydia trachomatis* protein GrgA activates transcription by contacting the nonconserved region of 66. *Proc Natl Acad Sci U S A* **109**: 16870–16875.

Barta, M.L., Hickey, J., Kemege, K.E., Lovell, S., Battaile, K.P., and Hefty, P.S. (2013) Structure of CT584 from *Chlamydia trachomatis* refined to 3.05 Å resolution. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **69**: 1196–1201.

Bastidas, R.J., Elwell, C.A., Engel, J.N., and Valdivia, R.H. (2013) Chlamydial intracellular survival strategies. *Cold Spring Harb Perspect Med* **3**: 1–21.

Bauler, L.D., and Hackstadt, T. (2014) Expression and Targeting of secreted proteins from *Chlamydia trachomatis*. *J Bacteriol* **196**: 1325–1334.

Beatty, W.L. (2006) Trafficking from CD63-positive late endocytic multivesicular bodies is essential for intracellular development of *Chlamydia trachomatis*. *J Cell Sci* **119**: 350–359.

Beatty, W.L. (2008) Late endocytic multivesicular bodies intersect the chlamydial inclusion in the absence of CD63. *Infect Immun* **76**: 2872–2881.

Bébéar, C.M., and Barbeyrac, B. de (2009) Genital *Chlamydia trachomatis* infections. *Clin Microbiol Infect* **15**: 4–10.

Becker, E., and Hegemann, J.H. (2014) All subtypes of the Pmp adhesin family are implicated in chlamydial virulence and show species-specific function. *Microbiologyopen* **3**: 544–56.

Belland, R.J., Scidmore, M.A., Dean, D., Hogan, D., Whitmire, W.M., McClarty, G., and Caldwell, H.D. (2001) *Chlamydia trachomatis* cytotoxicity associated with complete and partial cytotoxin genes. *Proc Natl Acad Sci U S A* **98**: 13984–13989.

Belland, R.J., Zhong, G., Dean, D., Hogan, D., Sturdevant, D.E., Sharma, J., *et al.* (2003) Genomic transcriptional profiling of the developmental cycle of *Chlamydia trachomatis*. *Proc Natl Acad Sci U S A* **100**: 8478–8483.

Betts, H.J., Wolf, K., and Fields, K.A. (2009) Effector protein modulation of host cells: examples in the *Chlamydia* spp. arsenal. *Curr Opin Microbiol* **12**: 81–87.

Binet, R., and Maurelli, A.T. (2005) Frequency of spontaneous mutations that confer antibiotic resistance in *Chlamydia* spp. *Antimicrob Agents Chemother* **49**: 2865–2873.

Binet, R., and Maurelli, A.T. (2009) Transformation and isolation of allelic exchange mutants of *Chlamydia psittaci* using recombinant DNA introduced by electroporation. *Proc Natl Acad Sci U S A* **106**: 292–297.

Blasi, F., Tarsia, P., and Aliberti, S. (2009) *Chlamydophila pneumoniae*. *Clin Microbiol Infect* **15**: 29–35.

Böcker, S., Heurich, A., Franke, C., Monajembashi, S., Sachse, K., Saluz, H.P., and Hänel, F. (2014) *Chlamydia psittaci* inclusion membrane protein IncB associates with host protein Snapin. *Int J Med Microbiol* **304**: 542–553.

Bodetti, T.J., Jacobson, E., Wan, C., Hafner, L., Pospischil, A., Rose, K., and Timms, P. (2002) Molecular evidence to support the expansion of the hostrange of *Chlamydophila pneumoniae* to include reptiles as well as humans, horses, koalas and amphibians. *Syst Appl Microbiol* **25**: 146–152.

Boncompain, G., Müller, C., Meas-Yedid, V., Schmitt-Kopplin, P., Lazarow, P.B., and Subtil, A. (2014) The intracellular bacteria *Chlamydia* hijack peroxisomes and utilize their enzymatic capacity to produce bacteria-specific phospholipids. *PLoS One* **9**.

Borges, V., Ferreira, R., Nunes, A., Nogueira, P.J., Borrego, M.J., and Gomes, J.P. (2010) Normalization strategies for real-time expression data in *Chlamydia trachomatis*. *J Microbiol Methods* **82**: 256–264.

Borges, V., Nunes, A., Ferreira, R., Borrego, M.J., and Gomes, J.P. (2012) Directional evolution of *Chlamydia trachomatis* towards niche-specific adaptation. *J Bacteriol* **194**: 6143–6153.

Bothe, M., Dutow, P., Pich, A., Genth, H., and Klos, A. (2015) DXD Motif-Dependent and -Independent Effects of the *Chlamydia trachomatis* Cytotoxin CT166. *Toxins (Basel)* **7**: 621–637.

Brinkworth, A.J., Malcolm, D.S., Pedrosa, A.T., Roguska, K., Shahbazian, S., Graham, J.E., *et al.* (2011) *Chlamydia trachomatis* Slc1 is a type III secretion chaperone that enhances the translocation of its invasion effector substrate TARP. *Mol Microbiol* **82**: 131–144.

Brown, H.M., Knowlton, A.E., Snavely, E., Nguyen, B.D., Richards, T.S., and Grieshaber, S.S. (2014) Multinucleation during *C. trachomatis* infections is caused by the contribution of two effector pathways. *PLoS One* **9**: 1–14.

Brunelle, B.W., Nicholson, T.L., and Stephens, R.S. (2004) Microarray-based genomic surveying of gene polymorphisms in *Chlamydia trachomatis*. *Genome Biol* **5**: R42.

Bullock, H.D., Hower, S., and Fields, K.A. (2012) Domain analyses reveal that *Chlamydia trachomatis* CT694 protein belongs to the membrane-localized family of type III effector proteins. *J Biol Chem* **287**: 28078–28086.

Cai, Y., Fukushi, H., Koyasu, S., Kuroda, E., Yamaguchi, T., and Hirai, K. (2002) An etiological investigation of domestic cats with conjunctivitis and upper respiratory tract disease in Japan. *J Vet Med Sci* **64**: 215–219.

Caldwell, H.D., Wood, H., Dean, D., Bailey, R., Jones, R.B., Mabey, D., *et al.* (2003) Polymorphisms in *Chlamydia trachomatis* tryptophan synthase genes differentiate between genital and ocular isolates. *J Clin Invest* **111**: 1757–1769.

Cao, B., Zhao, Y., Kou, Y., Ni, D., Zhang, X.C., and Huang, Y. (2014) Structure of the nonameric bacterial amyloid secretion channel. *Proc Natl Acad Sci U S A* **111**: E5439–44.

Capmany, A., and Damiani, M.T. (2010) *Chlamydia trachomatis* intercepts golgi-derived sphingolipids through a rab14-mediated transport required for bacterial development and replication. *PLoS One* **5**.

Carabeo, R.A., Dooley, C.A., Grieshaber, S.S., and Hackstadt, T. (2007) Rac interacts with Abi-1 and WAVE2 to promote an Arp2/3-dependent actin recruitment during chlamydial invasion. *Cell Microbiol* **9**: 2278–2288.

Carabeo, R.A., Grieshaber, S.S., Fischer, E.R., and Hackstadt, T. (2002) *Chlamydia trachomatis* Induces Remodeling of the Actin Cytoskeleton during Attachment and Entry into HeLa Cells. **70**: 3793–3803.

Carlson, J.H., Hughes, S., Hogan, D., Cieplak, G., Sturdevant, D.E., McClarty, G., *et al.* (2004) Polymorphisms in the *Chlamydia trachomatis* cytotoxin locus associated with ocular and genital isolates. *Infect Immun* **72**: 7063–7072.

Carlson, J.H., Porcella, S.F., Mcclarty, G., and Caldwell, H.D. (2005) Comparative Genomic Analysis of *Chlamydia trachomatis* Oculotropic and Genitotropic Strains. *Infect Immun* **73**: 6407–6418.

Carlson, J.H., Whitmire, W.M., Crane, D.D., Wicke, L., Virtaneva, K., Sturdevant, D.E., *et al.* (2008) The *Chlamydia trachomatis* plasmid is a transcriptional regulator of chromosomal genes and a virulence factor. *Infect Immun* **76**: 2273–2283.

Case, E.D.R., Peterson, E.M., and Tan, M. (2010) Promoters for *Chlamydia* type III secretion genes show a differential response to DNA supercoiling that correlates with temporal expression pattern. *J Bacteriol* **192**: 2569–2574.

Celli, J. (2015) The changing nature of the *Brucella*-containing vacuole. *Cell Microbiol* **17**: 951–958.

Chellas-Géry, B., Linton, C.N., and Fields, K.A. (2007) Human GCIP interacts with CT847, a novel *Chlamydia trachomatis* type III secretion substrate, and is degraded in a tissue-culture infection model. *Cell Microbiol* **9**: 2417–2430.

Chen, A.L., Johnson, K. a., Lee, J.K., Sütterlin, C., and Tan, M. (2012) CPAF: A Chlamydial Protease in Search of an Authentic Substrate. *PLoS Pathog* **8**: e1002842.

Chen, C., Chen, D., Sharma, J., Cheng, W., Zhong, Y., Liu, K., *et al.* (2006) The hypothetical protein CT813 is localized in the *Chlamydia trachomatis* inclusion membrane and is immunogenic in women urogenitally infected with C. trachomatis. *Infect Immun* **74**: 4826–4840.

Chen, Y.-S., Bastidas, R.J., Saka, H.A., Carpenter, V.K., Richards, K.L., Plano, G. V., and Valdivia, R.H. (2014) The *Chlamydia trachomatis* Type III Secretion Chaperone Slc1 Engages Multiple Early Effectors, Including TepP, a Tyrosine-phosphorylated Protein Required for the Recruitment of CrkI-II to Nascent Inclusions and Innate Immune Signaling. *PLoS Pathog* **10**.

Cheng, E., and Tan, M. (2012) Differential effects of DNA supercoiling on *Chlamydia* early promoters correlate with expression patterns in midcycle. *J Bacteriol* **194**: 3109–3115.

Chin, E., Kirker, K., Zuck, M., James, G., and Hybiske, K. (2012) Actin Recruitment to the *Chlamydia* Inclusion Is Spatiotemporally Regulated by a Mechanism That Requires Host and Bacterial Factors. *PLoS One* **7**: 1–12.

Chopra, I., Storey, C., Falla, T.J., and Pearce, J.H. (1998) Antibiotics, peptidoglycan synthesis and genomics: The chlamydial anomaly revisited. *Microbiology* **144**: 2673–2678.

Chumduri, C., Gurumurthy, R.K., Zadora, P.K., Mi, Y., and Meyer, T.F. (2013) *Chlamydia* infection promotes host DNA damage and proliferation but impairs the DNA damage response. *Cell Host Microbe* **13**: 746–758.

Clercq, E. De, Kalmar, I., and Vanrompay, D. (2013) Animal models for studying female genital tract infection with *Chlamydia trachomatis*. *Infect Immun* **81**: 3060–3067.

Clifton, D.R., Fields, K.A., Grieshaber, S.S., Dooley, C.A., Fischer, E.R., Mead, D.J., *et al.* (2004) A chlamydial type III translocated protein is tyrosine-phosphorylated at the site of entry and associated with recruitment of actin. *Proc Natl Acad Sci U S A* **101**: 10166–10171.

Cocchiaro, J., Kumar, Y., Fischer, E.R., Hackstadt, T., and Valdivia, R.H. (2008) Cytoplasmic lipid droplets are translocated into the lumen of the *Chlamydia trachomatis* parasitophorous vacuole. *Proc Natl Acad Sci U S A* **105**: 9379–9384.

Collingro, A., Tischler, P., Weinmaier, T., Penz, T., Heinz, E., Brunham, R., *et al.* (2011) Unity in Variety--The Pan-Genome of the *Chlamydiae*. *Mol Biol Evol* **28**: 3253–3270.

Cortes, C., Rzomp, K.A., Tvinnereim, A., Scidmore, M.A., and Wizel, B. (2007) *Chlamydia pneumoniae* inclusion membrane protein Cpn0585 interacts with multiple Rab GTPases. *Infect Immun* **75**: 5586–96.

Cossart, P., and Toledo-Arana, A. (2008) *Listeria monocytogenes*, a unique model in infection biology: an overview. *Microbes Infect* **10**: 1041–1050.

Costa, T.R.D., Felisberto-Rodrigues, C., Meir, A., Prevost, M.S., Redzej, A., Trokter, M., and Waksman, G. (2015) Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nat Rev Microbiol* **13**: 343–359.

Dahlström, L.A., Andersson, K., Luostarinen, T., Thoresen, S., Ögmundsdottír, H., Tryggvadottír, L., *et al.* (2011) Prospective seroepidemiologic study of human papillomavirus and other risk factors in cervical cancer. *Cancer Epidemiol Biomarkers Prev* **20**: 2541–2550.

Dehoux, P., Flores, R., Dauga, C., Zhong, G., and Subtil, A. (2011) Multi-genome identification and characterization of chlamydiae-specific type III secretion substrates: the Inc proteins. *BMC Genomics* **12**: 109.

Delevoye, C., Nilges, M., Dehoux, P., Paumet, F., Perrinet, S., Dautry-Varsat, A., and Subtil, A. (2008) SNARE protein mimicry by an intracellular bacterium. *PLoS Pathog* **4**.

DeMars, R., Weinfurter, J., Guex, E., Lin, J., and Potucek, Y. (2007) Lateral Gene

Transfer In Vitro in the Intracellular Pathogen *Chlamydia trachomatis*. *J Bacteriol* **189**: 991–1003.

Derré, I., Swiss, R., and Agaisse, H. (2011) The lipid transfer protein CERT interacts with the *Chlamydia* inclusion protein IncD and participates to ER-*Chlamydia* inclusion membrane contact sites. *PLoS Pathog* **7**.

Dinis, M., Cordeiro, D., Santo, I., Azevedo, J., Gomes, J.P., and Maria José Borrego (2015) Observações - Boletim Epidemiológico - Diagnóstico laboratorial da infeção por *Chlamydia trachomatis*, 1991-2014. .

Donati, M., Huot-creasy, H., Humphrys, M., Paolo, D., Francesco, D., and Myers, G. (2014) Genome Sequence of *Chlamydia suis* MD56 , Isolated from the Conjunctiva of a Weaned Piglet. **2**: 2147.

Dumoux, M., Clare, D.K., Saibil, H.R., and Hayward, R.D. (2012) *Chlamydiae* assemble a pathogen synapse to hijack the host endoplasmic reticulum. *Traffic* **13**: 1612–1627.

Dumoux, M., Menny, A., Delacour, D., and Hayward, R.D. (2015) A *Chlamydia* effector recruits CEP170 to reprogram host microtubule organization. *J Cell Sci* **44**.

ECDC (2014) Sexually transmitted infections in Europe 2012. .

Elwell, C. a., Jiang, S., Kim, J.H., Lee, A., Wittmann, T., Hanada, K., *et al.* (2011) *Chlamydia trachomatis* co-opts gbf1 and cert to acquire host sphingomyelin for distinct roles during intracellular development. *PLoS Pathog* **7**.

Engström, P., Bergström, M., Alfaro, A.C., Syam Krishnan, K., Bahnan, W., Almqvist, F., and Bergström, S. (2015) Expansion of the *Chlamydia trachomatis* inclusion does not require bacterial replication. *Int J Med Microbiol* .

Everett, K.D., Bush, R.M., and Andersen, A.A. (1999) Emended description of the order *Chlamydiales*, proposal of *Parachlamydiaceae* fam. nov. and *Simkaniaceae* fam. nov., each containing one monotypic genus, revised taxonomy of the family *Chlamydiaceae*, including a new genus and five new s. *Int J Syst Bacteriol* **49 Pt 2**: 415–40.

Fehlner-Gardiner, C., Roshick, C., Carlson, J.H., Hughes, S., Belland, R.J., Caldwell, H.D., and McClarty, G. (2002) Molecular basis defining human *Chlamydia trachomatis* tissue tropism: A possible role for tryptophan synthase. *J Biol Chem* **277**: 26893–26903.

Feldman, M.F., and Cornelis, G.R. (2003) The multitalented type III chaperones: All you can do with 15 kDa. *FEMS Microbiol Lett* **219**: 151–158.

Ferreira, R., Borges, V., Nunes, A., Borrego, M.J., and Gomes, J.P. (2013) Assessment of the load and transcriptional dynamics of *Chlamydia trachomatis* plasmid according to strains' tissue tropism. *Microbiol Res* **168**: 333–339.

Ferreira, R., Borges, V., Nunes, A., Nogueira, P.J., Borrego, M.J., and Gomes, J.P. (2012) Impact of Loci Nature on Estimating Recombination and Mutation Rates in *Chlamydia trachomatis*. *G3; Genes|Genomes|Genetics* **2**: 761–768.

Fields, K.A., Fischer, E.R., Mead, D.J., and Hackstadt, T. (2005) Analysis of Putative *Chlamydia trachomatis* Chaperones Scc2 and Scc3 and Their Use in the Identification of Type III Secretion Substrates. **187**: 6466–6478.

Fields, K.A., and Hackstadt, T. (2000) Evidence for the secretion of *Chlamydia trachomatis* CopN by a type III secretion mechanism. *Mol Microbiol* **38**: 1048–1060.

Fields, K.A., and Hackstadt, T. (2002) The chlamydial inclusion: escape from the endocytic pathway. *Annu Rev Cell Dev Biol* **18**: 221–245.

Fling, S.P., Sutherland, R. a, Steele, L.N., Hess, B., D'Orazio, S.E., Maisonneuve, J., *et*

al. (2001) CD8+ T cells recognize an inclusion membrane-associated protein from the vacuolar pathogen *Chlamydia trachomatis*. *Proc Natl Acad Sci U S A* **98**: 1160–1165.

Flores, R., and Zhong, G. (2015) The *Chlamydia pneumoniae* Inclusion Membrane Protein Cpn1027 Interacts with Host Cell Wnt Signaling Pathway Regulator Cytoplasmic Activation/Proliferation-Associated Protein 2 (Caprin2). *PLoS One* **10**: e0127909.

Furtado, A.R., Essid, M., Perrinet, S., Balañá, M.E., Yoder, N., Dehoux, P., and Subtil, A. (2013) The chlamydial OTU domain-containing protein ChlaOTU is an early type III secretion effector targeting ubiquitin and NDP52. *Cell Microbiol* **15**: 2064–2079.

Galán, J.E., Lara-Tejero, M., Marlovits, T.C., and Wagner, S. (2014) Bacterial Type III Secretion Systems: Specialized Nanomachines for Protein Delivery into Target Cells. *Annu Rev Microbiol* 415–438.

Gauliard, E., Ouellette, S.P., Rueden, K.J., and Ladant, D. (2015) Characterization of interactions between inclusion membrane proteins from *Chlamydia trachomatis*. *Front Cell Infect Microbiol* **5**: 1–11.

Gérard, H.C., Mishra, M.K., Mao, G., Wang, S., Hali, M., Whittum-Hudson, J. a., *et al.* (2013) Dendrimer-enabled DNA delivery and transformation of *Chlamydia pneumoniae*. *Nanomedicine Nanotechnology, Biol Med* **9**: 996–1008.

Giles, D.K., and Wyrick, P.B. (2008) Trafficking of chlamydial antigens to the endoplasmic reticulum of infected epithelial cells. *Microbes Infect* **10**: 1494–1503.

Gomes, J.P., Nunes, A., Bruno, W.J., Borrego, M.J., Florindo, C., and Dean, D. (2006) Polymorphisms in the nine polymorphic membrane proteins of *Chlamydia trachomatis* across all serovars: Evidence for serovar Da recombination and correlation with tissue tropism. *J Bacteriol* **188**: 275–286.

Gong, S., Lei, L., Chang, X., Belland, R.J., and Zhong, G. (2011) *Chlamydia trachomatis* secretion of hypothetical protein CT622 into host cell cytoplasm via a secretion pathway that can be inhibited by the type III secretion system inhibitor compound 1. *Microbiology* **157**: 1134–1144.

Gong, S., Yang, Z., Lei, L., Shen, L., and Zhong, G. (2013) Characterization of *Chlamydia trachomatis* plasmid-encoded open reading frames. *J Bacteriol* **195**: 3819–3826.

González, E., Rother, M., Kerr, M.C., Al-Zeer, M. a., Abu-Lubad, M., Kessler, M., *et al.* (2014) *Chlamydia* infection depends on a functional MDM2-p53 axis. *Nat Commun* **5**: 5201.

Greene, W., and Zhong, G. (2003) Inhibition of host cell cytokinesis by *Chlamydia trachomatis* infection. *J Infect* **47**: 45–51.

Grieshaber, S.S., Grieshaber, N. a., and Hackstadt, T. (2003) *Chlamydia trachomatis* uses host cell dynein to traffic to the microtubule-organizing center in a p50 dynamitin-independent process. *J Cell Sci* **116**: 3793–3802.

Grieshaber, S.S., Grieshaber, N. a., Miller, N.J., and Hackstadt, T. (2006) Chlamydia trachomatis causes centrosomal defects resulting in chromosomal segregation abnormalities. *Traffic* **7**: 940–949.

Griffiths, E., Ventresca, M.S., and Gupta, R.S. (2006) BLAST screening of chlamydial genomes to identify signature proteins that are unique for the *Chlamydiales*, *Chlamydiaceae*, *Chlamydophila* and *Chlamydia* groups of species. *BMC Genomics* **7**: 14.

Hackstadt, T., Rockey, D.D., Heinzen, R.A., and Scidmore, M.A. (1996) *Chlamydia*

trachomatis interrupts an exocytic pathway to acquire endogenously synthesized sphingomyelin in transit from the Golgi apparatus to the plasma membrane. *EMBO J* **15**: 964–977.

Hackstadt, T., Scidmore, M.A., Shaw, E.I., and Fischer, E.R. (1999) The *Chlamydia trachomatis* IncA protein is required for homotypic vesicle fusion. *Cell Microbiol* **1**: 119–130.

Harris, S.R., Clarke, I.N., Seth-Smith, H.M.B., Solomon, A.W., Cutcliffe, L.T., Marsh, P., *et al.* (2012) Whole-genome analysis of diverse *Chlamydia trachomatis* strains identifies phylogenetic relationships masked by current clinical typing. *Nat Genet* **44**: 413–419.

Hefty, P.S., and Stephens, R.S. (2007) Chlamydial type III secretion system is encoded on ten operons preceded by Sigma 70-like promoter elements. *J Bacteriol* **189**: 198–206.

Heinz, E., Rockey, D.D., Montanaro, J., Aistleitner, K., Wagner, M., and Horn, M. (2010) Inclusion membrane proteins of *Protochlamydia amoebophila* UWE25 reveal a conserved mechanism for host cell interaction among the *Chlamydiae*. *J Bacteriol* **192**: 5093–5102.

Heuer, D., Rejman Lipinski, A., Machuy, N., Karlas, A., Wehrens, A., Siedler, F., *et al.* (2009) *Chlamydia* causes fragmentation of the Golgi compartment to ensure reproduction. *Nature* **457**: 731–735.

Hilbi, H., and Haas, A. (2012) Secretive Bacterial Pathogens and the Secretory Pathway. *Traffic* **13**: 1187–1197.

Ho, T.D., and Starnbach, M.N. (2005) The *Salmonella enterica* serovar Typhimurium-encoded type III secretion systems can translocate *Chlamydia trachomatis* proteins into the cytosol of host cells. *Infect Immun* **73**: 905–911.

Hobolt-Pedersen, A.S., Christiansen, G., Timmerman, E., Gevaert, K., and Birkelund, S. (2009) Identification of *Chlamydia trachomatis* CT621, a protein delivered through the type III secretion system to the host cell cytoplasm and nucleus. *FEMS Immunol Med Microbiol* **57**: 46–58.

Horn, M. (2008) *Chlamydiae* as symbionts in eukaryotes. *Annu Rev Microbiol* **62**: 113–131.

Hovis, K.M., Mojica, S., Mcdermott, J.E., Pedersen, L., Simhi, C., Rank, R.G., *et al.* (2013) Genus-optimized strategy for the identification of chlamydial type III secretion substrates. *Pathog Dis* **69**: 213–222.

Hower, S., Wolf, K., and Fields, K.A. (2009) Evidence that CT694 is a novel *Chlamydia trachomatis* T3S substrate capable of functioning during invasion or early cycle development. *Mol Microbiol* **72**: 1423–1437.

Hybiske, K., and Stephens, R.S. (2007) Mechanisms of host cell exit by the intracellular bacterium *Chlamydia*. *Proc Natl Acad Sci U S A* **104**: 11430–11435.

Jahn, R., and Scheller, R.H. (2006) SNAREs - engines for membrane fusion. *Nat Rev Mol cell Biol* **7**: 631–643.

Jeffrey, B.M., Suchland, R.J., Quinn, K.L., Davidson, J.R., Stamm, W.E., and Rockey, D.D. (2010) Genome sequencing of recent clinical *Chlamydia trachomatis* strains identifies loci associated with tissue tropism and regions of apparent recombination. *Infect Immun* **78**: 2544–2553.

Jewett, T.J., Fischer, E.R., Mead, D.J., and Hackstadt, T. (2006) Chlamydial TARP is a bacterial nucleator of actin. *Proc Natl Acad Sci U S A* **103**: 15599–15604.

Jewett, T.J., Miller, N.J., Dooley, C.A., and Hackstadt, T. (2010) The conserved tarp actin binding domain is important for chlamydial invasion. *PLoS Pathog* **6**: 1–11.

Johnson, C.M., and Fisher, D.J. (2013) Site-specific, insertional inactivation of *incA* in *Chlamydia trachomatis* using a group II intron. *PLoS One* **8**.

Johnson, K. a., Tan, M., and Sütterlin, C. (2009) Centrosome abnormalities during a *Chlamydia trachomatis* infection are caused by dysregulation of the normal duplication pathway. *Cell Microbiol* **11**: 1064–1073.

Joseph, S.J., Didelot, X., Gandhi, K., Dean, D., and Read, T.D. (2011) Interplay of recombination and selection in the genomes of *Chlamydia trachomatis*. *Biol Direct* **6**: 28.

Kabeiseman, E.J., Cichos, K., Hackstadt, T., Lucas, A., and Moore, E.R. (2013) Vesicle-associated membrane protein 4 and syntaxin 6 interactions at the chlamydial inclusion. *Infect Immun* **81**: 3326–3337.

Kadota, K., Ishino, T., Matsuyama, T., Chinzei, Y., and Yuda, M. (2004) Essential role of membrane-attack protein in malarial transmission to mosquito host. *Proc Natl Acad Sci U S A* **101**: 16310–16315.

Kafsack, B.F.C., Pena, J.D.O., Coppens, I., Ravindran, S., Boothroyd, J.C., and Carruthers, V.B. (2009) Rapid membrane disruption by a perforin-like protein facilitates parasite exit from host cells. *Science* **323**: 530–533.

Kari, L., Goheen, M.M., Randall, L.B., Taylor, L.D., Carlson, J.H., Whitmire, W.M., *et al.* (2011) Generation of targeted *Chlamydia trachomatis* null mutants. *Proc Natl Acad Sci U S A* **108**: 7189–7193.

Kari, L., Southern, T.R., Downey, C.J., Watkins, H.S., Randall, L.B., Taylor, L.D., *et al.* (2014) *Chlamydia trachomatis* polymorphic membrane protein D is a virulence factor involved in early host-cell interactions. *Infect Immun* **82**: 2756–2762.

Karunakaran, K., Subbarayal, P., Vollmuth, N., and Rudel, T. (2015) *Chlamydia*-infected cells shed Gp96 to prevent chlamydial re-infection. *Mol Microbiol* .

Kern, J.M., Maass, V., and Maass, M. (2009) Molecular pathogenesis of chronic *Chlamydia pneumoniae* infection: a brief overview. *Clin Microbiol Infect* **15**: 36–41.

Kim, J.H., Jiang, S., Elwell, C. a., and Engel, J.N. (2011) *Chlamydia trachomatis* co-opts the FGF2 signaling pathway to enhance infection. *PLoS Pathog* **7**.

Kleba, B., and Stephens, R.S. (2008) Chlamydial effector proteins localized to the host cell cytoplasmic compartment. *Infect Immun* **76**: 4842–4850.

Knittler, M.R., Berndt, A., Böcker, S., Dutow, P., Hänel, F., Heuer, D., *et al.* (2014) *Chlamydia psittaci*: New insights into genomic diversity, clinical pathology, host-pathogen interaction and anti-bacterial immunity. *Int J Med Microbiol* **304**: 877–893.

Knowlton, A.E., Brown, H.M., Richards, T.S., Andreolas, L. a., Patel, R.K., and Grieshaber, S.S. (2011) *Chlamydia trachomatis* infection causes mitotic spindle pole defects independently from its effects on centrosome amplification. *Traffic* **12**: 854–866.

Kokes, M., Dunn, J.D., Granek, J.A., Nguyen, B.D., Barker, J.R., Valdivia, R.H., and Bastidas, R.J. (2015) Integrating Chemical Mutagenesis and Whole-Genome Sequencing as a Platform for Forward and Reverse Genetic Analysis of *Chlamydia*. *Cell Host Microbe* 1–10.

Kong, F.Y.S., and Hocking, J.S. (2015) Treatment challenges for urogenital and anorectal *Chlamydia trachomatis*. *BMC Infect Dis* **15**: 293.

Koskela, P., Anttila, T., Bjørge, T., Brunsvig, A., Dillner, J., Hakama, M., *et al.* (2000)

Chlamydia trachomatis infection as a risk factor for invasive cervical cancer. *Int J Cancer* **85**: 35–39.

Kumar, Y., Cocchiaro, J., and Valdivia, R.H. (2006) The Obligate Intracellular Pathogen *Chlamydia trachomatis* Targets Host Lipid Droplets. *Curr Biol* **16**: 1646–1651.

Kumar, Y., and Valdivia, R.H. (2008) Actin and Intermediate Filaments Stabilize the *Chlamydia trachomatis* Vacuole by Forming Dynamic Structural Scaffolds. *Cell Host Microbe* **4**: 159–169.

Kumar, Y., and Valdivia, R.H. (2009) Leading a Sheltered Life: Intracellular Pathogens and Maintenance of Vacuolar Compartments. *Cell Host Microbe* **5**: 593–601.

Lad, S.P., Li, J., Silva Correia, J. da, Pan, Q., Gadwal, S., Ulevitch, R.J., and Li, E. (2007) Cleavage of p65/RelA of the NF-kappaB pathway by *Chlamydia*. *Proc Natl Acad Sci U S A* **104**: 2933–2938.

Lad, S.P., Yang, G., Scott, D. a., Wang, G., Nair, P., Mathison, J., *et al.* (2007) Chlamydial CT441 Is a PDZ domain-containing tail-specific protease that interferes with the NF-kappaB pathway of immune response. *J Bacteriol* **189**: 6619–6625.

Lamoth, F., Pillonel, T., and Greub, G. (2015) *Waddlia*: an Emerging Pathogen and a Model Organism to Study the Biology of *Chlamydiae*. *Microbes Infect* 1–6.

Lane, B.J., Mutchler, C., Khodor, S. Al, Grieshaber, S.S., and Carabeo, R.A. (2008) Chlamydial entry involves TARP binding of guanine nucleotide exchange factors. *PLoS Pathog* **4**.

LaRock, D.L., Chaudhary, A., and Miller, S.I. (2015) *Salmonellae* interactions with host processes. *Nat Rev Microbiol* **13**: 191–205.

Lei, L., Dong, X., Li, Z., and Zhong, G. (2013) Identification of a Novel Nuclear Localization Signal Sequence in *Chlamydia trachomatis*-Secreted Hypothetical Protein CT311. *PLoS One* **8**: 1–7.

Lei, L., Qi, M., Budrys, N., Schenken, R., and Zhong, G. (2011) Localization of *Chlamydia trachomatis* hypothetical protein CT311 in host cell cytoplasm. *Microb Pathog* **51**: 101–109.

Leiva, N., Capmany, A., and Damiani, M.T. (2013) Rab11-Family of Interacting Protein 2 associates with chlamydial inclusions through its Rab-binding domain and promotes bacterial multiplication. *Cell Microbiol* **15**: 114–129.

Li, Z., Chen, C., Chen, D., Wu, Y., Zhong, Y., and Zhong, G. (2008) Characterization of fifty putative inclusion membrane proteins encoded in the *Chlamydia trachomatis* genome. *Infect Immun* **76**: 2746–2757.

Li, Z., Chen, D., Zhong, Y., Wang, S., and Zhong, G. (2008) The chlamydial plasmid-encoded protein pgp3 is secreted into the cytosol of *Chlamydia*-infected cells. *Infect Immun* **76**: 3415–3428.

Liechti, G.W., Kuru, E., Hall, E., Kalinda, A., Brun, Y. V, VanNieuwenhze, M., and Maurelli, a T. (2014) A new metabolic cell-wall labelling method reveals peptidoglycan in *Chlamydia trachomatis*. *Nature* **506**: 507–10.

Longbottom, D., and Coulter, L.J. (2003) Animal chlamydioses and zoonotic implications. *J Comp Pathol* **128**: 217–244.

Löwer, M., and Schneider, G. (2009) Prediction of type III secretion signals in genomes of gram-negative bacteria. *PLoS One* **4**: 1435–1439.

Lu, C., Lei, L., Peng, B., Tang, L., Ding, H., Gong, S., *et al.* (2013) *Chlamydia*

trachomatis GlgA Is Secreted into Host Cell Cytoplasm. *PLoS One* **8**: 1–12.

Lutter, E.I., Barger, A.C., Nair, V., and Hackstadt, T. (2013) *Chlamydia trachomatis* Inclusion Membrane Protein CT228 Recruits Elements of the Myosin Phosphatase Pathway to Regulate Release Mechanisms. *Cell Rep* **3**: 1921–1931.

Lutter, E.I., Bonner, C., Holland, M.J., Suchland, R.J., Stamm, W.E., Jewett, T.J., *et al.* (2010) Phylogenetic analysis of *Chlamydia trachomatis* tarp and correlation with clinical phenotype. *Infect Immun* **78**: 3678–3688.

Lutter, E.I., Martens, C., and Hackstadt, T. (2012) Evolution and conservation of predicted inclusion membrane proteins in *Chlamydiae*. *Comp Funct Genomics* **2012**.

Mallios, R.R., Ojcius, D.M., and Ardell, D.H. (2009) An iterative strategy combining biophysical criteria and duration hidden Markov models for structural predictions of *Chlamydia trachomatis* sigma66 promoters. *BMC Bioinformatics* **10**: 271.

Markham, A.P., Jaafar, Z.A., Kemege, K.E., Middaugh, C.R., and Hefty, P.S. (2009) Biophysical characterization of *Chlamydia trachomatis* CT584 supports its potential role as a type III secretion needle tip protein. *Biochemistry* **48**: 10353–10361.

Mathews, S., and Timms, P. (2006) *In silico* identification of chlamydial promoters and their role in regulation of development. In *Chlamydia: genomics and pathogenesis*. Bavoil, P.M., and Wyrick, P.B. (eds). Horizon Bioscience, Norfolk, United Kingdom. pp. 133–156.

McBride, M.J., and Nakane, D. (2015) Flavobacterium gliding motility and the type IX secretion system. *Curr Opin Microbiol* **28**: 72–77.

Mendonça, A.G., Alves, R.J., and Pereira-Leal, J.B. (2011) Loss of genetic redundancy in reductive genome evolution. *PLoS Comput Biol* **7**: e1001082.

Mirrashidi, K.M., Elwell, C.A., Verschueren, E., Johnson, J.R., Frando, A., Dollen, J. Von, *et al.* (2015) Global Mapping of the Inc-Human Interactome Reveals that Retromer Restricts *Chlamydia* Infection. *Cell Host Microbe* **18**: 109–121.

Misaghi, S., Balsara, Z.R., Catic, A., Spooner, E., Ploegh, H.L., and Starnbach, M.N. (2006) *Chlamydia trachomatis*-derived deubiquitinating enzymes in mammalian cells during infection. *Mol Microbiol* **61**: 142–150.

Mital, J., Lutter, E.I., Barger, A.C., Dooley, C.A., and Hackstadt, T. (2015) *Chlamydia trachomatis* inclusion membrane protein CT850 interacts with the dynein light chain DYNLT1 (Tctex1). *Biochem Biophys Res Commun* **1**: 1–6.

Mital, J., Miller, N.J., Dorward, D.W., Dooley, C.A., and Hackstadt, T. (2013) Role for Chlamydial Inclusion Membrane Proteins in Inclusion Membrane Structure and Biogenesis. *PLoS One* **8**.

Mital, J., Miller, N.J., Fischer, E.R., and Hackstadt, T. (2010) Specific chlamydial inclusion membrane proteins associate with active Src family kinases in microdomains that interact with the host microtubule network. *Cell Microbiol* **12**: 1235–1249.

Mojica, S., Hovis, K.M., Frieman, M.B., Tran, B., Hsia, R., Ravel, J., *et al.* (2015) SINC, a type III secreted protein of *Chlamydia psittaci*, targets the inner nuclear membrane of infected cells and uninfected neighbors. *Mol Biol Cell* .

Mölleken, K., Schmidt, E., and Hegemann, J.H. (2010) Members of the Pmp protein family of *Chlamydia pneumoniae* mediate adhesion to human cells via short repetitive peptide motifs. *Mol Microbiol* **78**: 1004–17.

Moorhead, A.M., Jung, J.Y., Smirnov, A., Kaufer, S., and Scidmore, M.A. (2010) Multiple

host proteins that function in phosphatidylinositol-4-phosphate metabolism are recruited to the chlamydial inclusion. *Infect Immun* **78**: 1990–2007.

Mota, L.J., and Cornelis, G.R. (2005) The bacterial injection kit : Type III secretion systems. *Ann Med* **4**: 234–239.

Mueller, C.A., Broz, P., and Cornelis, G.R. (2008) The type III secretion system tip complex and translocon. *Mol Microbiol* **68**: 1085–1095.

Mueller, K.E., and Fields, K.A. (2015) Application of β-Lactamase Reporter Fusions as an Indicator of Effector Protein Secretion during Infections with the Obligate Intracellular Pathogen *Chlamydia trachomatis*. *PLoS One* **10**: e0135295.

Mueller, K.E., Wolf, K., and Fields, K.A. (2016) Gene Deletion by Fluorescence-Reported Allelic Exchange Mutagenesis in *Chlamydia trachomatis*. *MBio* **7**: 1–9.

Muschiol, S., Boncompain, G., Vromman, F., Dehoux, P., Normark, S., Henriques-Normark, B., and Subtil, A. (2011) Identification of a family of effectors secreted by the type III secretion system that are conserved in pathogenic *Chlamydiae*. *Infect Immun* **79**: 571–580.

Nam, H.-J., Naylor, R.M., and Deursen, J.M. van (2015) Centrosome dynamics as a source of chromosomal instability. *Trends Cell Biol* **25**: 65–73.

Nans, A., Ford, C., and Hayward, R.D. (2015) Host-pathogen reorganisation during host cell entry by *Chlamydia trachomatis*. *Microbes Infect* 2–6.

Negrate, G. Le, Krieg, A., Faustin, B., Loeffler, M., Godzik, A., Krajewski, S., and Reed, J.C. (2008) ChlaDub1 of *Chlamydia trachomatis* suppresses NF-kappaB activation and inhibits IkappaBalpha ubiquitination and degradation. *Cell Microbiol* **10**: 1879–1892.

Nelson, D.E., Dean, D., Taylor, L.D., Dorward, D.W., Goheen, M.M., and Caldwell, H.D. (2006) Inhibition of *Chlamydiae* by primary alcohols correlates with the strain-specific complement of plasticity zone phospholipase D genes. *Infect Immun* **74**: 73–80.

Neuendorf, E., Gajer, P., Bowlin, A.K., Marques, P.X., Ma, B., Yang, H., *et al.* (2015) *Chlamydia caviae* infection alters abundance but not composition of the guinea pig vaginal microbiota. *Pathog Dis* 1–12.

Nguyen, B.D., Cunningham, D., Liang, X., Chen, X., Toone, E.J., Raetz, C.R.H., *et al.* (2011) Lipooligosaccharide is required for the generation of infectious elementary bodies in *Chlamydia trachomatis*. *Proc Natl Acad Sci U S A* **108**: 10284–10289.

Nguyen, B.D., and Valdivia, R.H. (2012) Virulence determinants in the obligate intracellular pathogen *Chlamydia trachomatis* revealed by forward genetic approaches. *Proc Natl Acad Sci U S A* **109**: 1263–1268.

Nicholson, T.L., Olinger, L., Chong, K., Schoolnik, G., and Stephens, R.S. (2003) Global stage-specific gene regulation during the developmental cycle of *Chlamydia trachomatis*. *J Bacteriol* **185**: 3179–3189.

Niehus, E., Cheng, E., and Tan, M. (2008) DNA Supercoiling-Dependent Gene Regulation in *Chlamydia*. *J Bacteriol* **190**: 6419–6427.

Nunes, A., Nogueira, P.J., Borrego, M.J., and Gomes, J.P. (2008) *Chlamydia trachomatis* diversity viewed as a tissue-specific coevolutionary arms race. *Genome Biol* **9**: R153.

O'Connell, C.M., Ingalls, R.R., Andrews Jr., C.W., Scurlock, A.M., and Darville, T. (2007) Plasmid-deficient *Chlamydia muridarum* fail to induce immune pathology and protect against oviduct disease. *J Immunol* **179**: 4027–4034.

Olive, A.J., Haff, M.G., Emanuele, M.J., Sack, L.M., Barker, J.R., Elledge, S.J., and Starnbach, M.N. (2014) *Chlamydia trachomatis*-induced alterations in the host cell proteome are required for intracellular growth. *Cell Host Microbe* **15**: 113–124.

Omsland, A., Sager, J., Nair, V., Sturdevant, D.E., and Hackstadt, T. (2013) Developmental stage-specific metabolic and transcriptional activity of *Chlamydia trachomatis* in an axenic medium. *Proc Natl Acad Sci U S A* **110**: 1970–1970.

Packiam, M., Weinrick, B., Jacobs, W.R., and Maurelli, A.T. (2015) Structural characterization of muropeptides from *Chlamydia trachomatis* peptidoglycan by mass spectrometry resolves “chlamydial anomaly.” *Proc Natl Acad Sci U S A* **112**: 11660–11665.

Page, A.L., and Parsot, C. (2002) Chaperones of the type III secretion pathway: Jacks of all trades. *Mol Microbiol* **46**: 1–11.

Pais, S. V., Milho, C., Almeida, F., and Mota, L.J. (2013) Identification of Novel Type III Secretion Chaperone-Substrate Complexes of *Chlamydia trachomatis*. *PLoS One* **8**.

Pennini, M.E., Perrinet, S., Dautry-Varsat, A., and Subtil, A. (2010) Histone methylation by NUE, a novel nuclear effector of the intracellular pathogen *Chlamydia trachomatis*. *PLoS Pathog* **6**: 1–12.

Peters, J., Wilson, D.P., Myers, G., Timms, P., and Bavoil, P.M. (2007) Type III secretion à la *Chlamydia*. *Trends Microbiol* **15**: 241–251.

Pokrovskaya, I.D., Szwedo, J.W., Goodwin, A., Lupashina, T. V., Nagarajan, U.M., and Lupashin, V. V. (2012) *Chlamydia trachomatis* hijacks intra-Golgi COG complex-dependent vesicle trafficking pathway. *Cell Microbiol* **14**: 656–668.

Polkinghorne, A., Hanger, J., and Timms, P. (2013) Recent advances in understanding the biology, epidemiology and control of chlamydial infections in koalas. *Vet Microbiol* **165**: 214–223.

Porcella, S.F., Carlson, J.H., Sturdevant, D.E., Sturdevant, G.L., Kanakabandi, K., Virtaneva, K., *et al.* (2015) Transcriptional Profiling of Human Epithelial Cells Infected with Plasmid-Bearing and Plasmid-Deficient *Chlamydia trachomatis*. *Infect Immun* **83**: 534–543.

Qi, M., Lei, L., Gong, S., Liu, Q., DeLisa, M.P., and Zhong, G. (2011) *Chlamydia trachomatis* secretion of an immunodominant hypothetical protein (CT795) into host cell cytoplasm. *J Bacteriol* **193**: 2498–2509.

Ramsey, K.H., Sigar, I.M., Schripsema, J.H., Denman, C.J., Bowlin, A.K., Myers, G., and Rank, R.G. (2009) Strain and virulence diversity in the mouse pathogen *Chlamydia muridarum*. *Infect Immun* **77**: 3284–3293.

Read, T.D., Joseph, S.J., Didelot, X., Liang, B., Patel, L., and Dean, D. (2013) Comparative analysis of *Chlamydia psittaci* genomes reveals the recent emergence of a pathogenic lineage with a broad host range. *MBio* **4**: 1–10.

Read, T.D., Myers, G., Brunham, R., Nelson, W.C., Paulsen, I.T., Heidelberg, J., *et al.* (2003) Genome sequence of *Chlamydophila caviae* (*Chlamydia psittaci* GPIC): Examining the role of niche-specific genes in the evolution of the *Chlamydiaceae*. *Nucleic Acids Res* **31**: 2134–2147.

Reinhold, P., Sachse, K., and Kaltenboeck, B. (2011) *Chlamydiaceae* in cattle: Commensals, trigger organisms, or pathogens? *Vet J* **189**: 257–267.

Richards, T.S., Knowlton, A.E., and Grieshaber, S.S. (2013) *Chlamydia trachomatis* homotypic inclusion fusion is promoted by host microtubule trafficking. *BMC Microbiol*

13: 185.

Rockey, D.D., Scidmore, M.A., Bannantine, J.P., and Brown, W.J. (2002) Proteins in the chlamydial inclusion membrane. *Microbes Infect* **4**: 333–340.

Ronzone, E., and Paumet, F. (2013) Two Coiled-Coil Domains of *Chlamydia trachomatis* IncA Affect Membrane Fusion Events during Infection. *PLoS One* **8**.

Ronzone, E., Wesolowski, J., Bauler, L.D., Bhardwaj, A., Hackstadt, T., and Paumet, F. (2014) An alfa-Helical Core Encodes the Dual Functions of the Chlamydial Protein IncA. *J Biol Chem* **289**: 33469–33480.

Rosado, C.J., Kondos, S., Bull, T.E., Kuiper, M.J., Law, R.H.P., Buckle, A.M., *et al.* (2008) The MACPF/CDC family of pore-forming toxins. *Cell Microbiol* **10**: 1765–1774.

Rosario, C.J., and Tan, M. (2012) The early gene product EUO is a transcriptional repressor that selectively regulates promoters of *Chlamydia* late genes. *Mol Microbiol* **84**: 1097–1107.

Rosmarin, D.M., Carette, J.E., Olive, A.J., Starnbach, M.N., Brummelkamp, T.R., and Ploegh, H.L. (2012) Attachment of *Chlamydia trachomatis* L2 to host cells requires sulfation. *Proc Natl Acad Sci U S A* **109**: 10059–10064.

Roulis, E., Polkinghorne, A., and Timms, P. (2013) *Chlamydia pneumoniae*: Modern insights into an ancient pathogen. *Trends Microbiol* **21**: 120–128.

Rzomp, K.A., Moorhead, A.R., and Scidmore, M.A. (2006) The GTPase Rab4 interacts with *Chlamydia trachomatis* inclusion membrane protein CT229. *Infect Immun* **74**: 5362–5373.

Rzomp, K.A., Scholtes, L.D., Briggs, B.J., Whittaker, G.R., and Scidmore, M.A. (2003) Rab GTPases are recruited to chlamydial inclusions in both a species-dependent and species-independent manner. *Infect Immun* **71**: 5855–5870.

Sachse, K., Laroucau, K., Riege, K., Wehner, S., Dilcher, M., Creasy, H.H., *et al.* (2014) Evidence for the existence of two new members of the family *Chlamydiaceae* and proposal of *Chlamydia avium* sp. nov. and *Chlamydia gallinacea* sp. nov. *Syst Appl Microbiol* **37**: 79–88 http://dx.doi.org/10.1016/j.syapm.2013.12.004.

Saka, H.A., Thompson, J.W., Chen, Y.-S., Dubois, L.G., Haas, J.T., Moseley, A., and Valdivia, R.H. (2015) *Chlamydia trachomatis* Infection Leads to Defined Alterations to the Lipid Droplet Proteome in Epithelial Cells. *PLoS One* **10**: e0124630.

Saka, H.A., Thompson, J.W., Chen, Y.-S., Kumar, Y., Dubois, L.G., Moseley, A., and Valdivia, R.H. (2011) Quantitative proteomics reveals metabolic and pathogenic properties of *Chlamydia trachomatis* developmental forms. *Mol Microbiol* **82**: 1185–1203.

Samudrala, R., Heffron, F., and Mcdermott, J.E. (2009) Accurate Prediction of Secreted Substrates and Identification of a Conserved Putative Secretion Signal for Type III Secretion Systems. *PLoS Pathog* **5**.

Schachter, J. (1999) Infection and disease epidemiology. In *Chlamydia: intracellular biology, pathogenesis and immunity*. Stephens, R.S. (ed.). ASM Press, Washington DC. pp. 139–169.

Schaik, E.J. van, Chen, C., Mertens, K., Weber, M.M., and Samuel, J.E. (2013) Molecular pathogenesis of the obligate intracellular bacterium *Coxiella burnetii*. *Nat Rev Microbiol* **11**: 561–573.

Schautteet, K., and Vanrompay, D. (2011) *Chlamydiaceae* infections in pig. *Vet Res* **42**: 29.

Schroeder, G.N., and Hilbi, H. (2008) Molecular Pathogenesis of *Shigella* spp.: Controlling Host Cell Signaling, Invasion, and Death by Type III Secretion. *Clin Microbiol Rev* **21**: 134–156.

Scidmore, M.A. (2005) Culture and Cultivation of *Chlamydia trachomatis*. *Curr Protoc Microbiol* 1–25.

Scidmore, M.A., Fischer, E.R., and Hackstadt, T. (2003) Restricted fusion of *Chlamydia trachomatis* vesicles with endocytic compartments during the initial stages of infection. *Infect Immun* **71**: 973–984.

Scidmore, M.A., and Hackstadt, T. (2001) Mammalian 14-3-3-beta associates with the *Chlamydia trachomatis* inclusion membrane via its interaction with IncG. *Mol Microbiol* **39**: 1638–1650.

Scidmore, M.A., Rockey, D.D., Fischer, E.R., Heinzen, R.A., and Hackstadt, T. (1996) Vesicular interactions of the *Chlamydia trachomatis* inclusion are determined by chlamydial early protein synthesis rather than route of entry. *Infect Immun* **64**: 5366–5372.

Seth-Smith, H.M.B., Harris, S.R., Persson, K., Marsh, P., Barron, A., Bignell, A., *et al.* (2009) Co-evolution of genomes and plasmids within *Chlamydia trachomatis* and the emergence in Sweden of a new variant strain. *BMC Genomics* **10**: 239.

Sharma, M., Machuy, N., Böhme, L., Karunakaran, K., Mäurer, A.P., Meyer, T.F., and Rudel, T. (2011) HIF-1α is involved in mediating apoptosis resistance to *Chlamydia trachomatis*-infected cells. *Cell Microbiol* **13**: 1573–1585.

Sharma, M., and Rudel, T. (2009) Apoptosis resistance in *Chlamydia*-infected cells: a fate worse than death? *FEMS Immunol Med Microbiol* **55**: 154–161.

Shaw, A.C., Christiansen, G., Roepstorff, P., and Birkelund, S. (2000) Genetic differences in the *Chlamydia trachomatis* tryptophan synthase alfa-subunit can explain variations in serovar pathogenesis. *Microbes Infect* **2**: 581–592.

Shaw, E.I., Dooley, C.A., Fischer, E.R., Scidmore, M.A., Fields, K.A., and Hackstadt, T. (2000) Three temporal classes of gene expression during the *Chlamydia trachomatis* developmental cycle. *Mol Microbiol* **37**: 913–925.

Sherwood, R.K., and Roy, C.R. (2013) A rab-centric perspective of bacterial pathogen-occupied vacuoles. *Cell Host Microbe* **14**: 256–268.

Shima, K., Klinger, M., Schütze, S., Kaufhold, I., Solbach, W., Reiling, N., and Rupp, J. (2015) The role of endoplasmic reticulum-related BiP/GRP78 in interferon gamma-induced persistent *Chlamydia pneumoniae* infection. *Cell Microbiol* **49**.

Silva-Herzog, E., Joseph, S.S., Avery, A.K., Coba, J. a., Wolf, K., Fields, K.A., and Plano, G. V. (2011) Scc1 (CP0432) and Scc4 (CP0033) function as a type III secretion chaperone for CopN of *Chlamydia pneumoniae*. *J Bacteriol* **193**: 3490–3496.

Simeone, R., Bottai, D., and Brosch, R. (2009) ESX/type VII secretion systems and their role in host-pathogen interaction. *Curr Opin Microbiol* **12**: 4–10.

Sisko, J.L., Spaeth, K.E., Kumar, Y., and Valdivia, R.H. (2006) Multifunctional analysis of *Chlamydia*-specific genes in a yeast expression system. *Mol Microbiol* **60**: 51–66.

Sixt, B.S., Siegl, A., Müller, C., Watzka, M., Wultsch, A., Tziotis, D., *et al.* (2013) Metabolic Features of *Protochlamydia amoebophila* Elementary Bodies - A Link between Activity and Infectivity in Chlamydiae. *PLoS Pathog* **9**.

Skipp, P., Robinson, J., O'Connor, C.D., and Clarke, I.N. (2005) Shotgun proteomic

analysis of *Chlamydia trachomatis*. *Proteomics* **5**: 1558–1573.

Snavely, E., Kokes, M., Dunn, J.D., Saka, H.A., Nguyen, B.D., Bastidas, R.J., *et al.* (2014) Reassessing the role of the secreted protease CPAF in *Chlamydia trachomatis* infection through genetic approaches. *Pathog Dis* **71**: 336–351.

Somani, J., Bhullar, V., Workowski, K., Farshy, C., and Black, C. (2000) Multiple drug-resistant *Chlamydia trachomatis* associated with clinical treatment failure. *J Infect Dis* **181**: 1421–1427.

Somboonna, N., Wan, R., Ojcius, D.M., Pettengill, M. a, Joseph, S.J., Chang, A., *et al.* (2011) Hypervirulent *Chlamydia trachomatis* Clinical Strain Is a Recombinant between Lymphogranuloma Venereum (L2) and D Lineages. *MBio* **2**: 1–12.

Song, L., Carlson, J.H., Whitmire, W.M., Kari, L., Virtaneva, K., Sturdevant, D.E., *et al.* (2013) *Chlamydia trachomatis* plasmid-encoded pgp4 is a transcriptional regulator of virulence-associated genes. *Infect Immun* **81**: 636–644.

Spaeth, K.E., Chen, Y.-S., and Valdivia, R.H. (2009) The *Chlamydia* type III secretion system C-ring engages a chaperone-effector protein complex. *PLoS Pathog* **5**.

Stary, G., Olive, A., Radovic-Moreno, a. F., Gondek, D., Alvarez, D., Basto, P. a., *et al.* (2015) A mucosal vaccine against *Chlamydia trachomatis* generates two waves of protective memory T cells. *Science* **348**: aaa8205–1 aaa8205–14.

Stephens, R.S., Kalman, S., Lammel, C., Fan, J., Marathe, R., Aravind, L., *et al.* (1998) Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*. *Science* **282**: 754–759.

Stone, C.B., Sugiman-Marangos, S., Bulir, D.C., Clayden, R.C., Leighton, T.L., Slootstra, J.W., *et al.* (2012) Structural characterization of a novel *Chlamydia pneumoniae* type III secretion-associated protein, Cpn0803. *PLoS One* **7**: e30220.

Subbarayal, P., Karunakaran, K., Winkler, A.-C., Rother, M., González, E., Meyer, T.F., and Rudel, T. (2015) EphrinA2 Receptor (EphA2) Is an Invasion and Intracellular Signaling Receptor for *Chlamydia trachomatis*. *PLOS Pathog* **11**: e1004846.

Subtil, A., Delevoye, C., Balañá, M.E., Tastevin, L., Perrinet, S., and Dautry-Varsat, A. (2005) A directed screen for chlamydial proteins secreted by a type III mechanism identifies a translocated protein and numerous other new candidates. *Mol Microbiol* **56**: 1636–1647.

Subtil, A., Parsot, C., and Dautry-Varsat, A. (2001) Secretion of predicted Inc proteins of *Chlamydia pneumoniae* by a heterologous type III machinery. *Mol Microbiol* **39**: 792–800.

Sun, H.S., Sin, A.T.-W., Poirier, M., and Harrison, R.E. (2015) *Chlamydia trachomatis* Inclusion Disrupts Host Cell Cytokinesis to Enhance Its Growth in Multinuclear Cells. *J Cell Biochem* .

Tan, C., Spitznagel, J.K., Shou, H., Hsia, R., and Bavoil, P.M. (2006) The polymorphic membrane protein gene family of the *Chlamydiaceae*. In *Chlamydia: Genomics and Pathogenesis*. Bavoil, P.M., and Wyrick, P.B. (eds). Horizon Bioscience, Norfolk, United Kingdom. pp. 195–218.

Tan, M., Gaal, T., Gourse, R.L., and Engel, J.N. (1998) Mutational analysis of the *Chlamydia trachomatis* rRNA P1 promoter defines four regions important for transcription in vitro. *J Bacteriol* **180**: 2359–2366.

Taylor, L.D., Nelson, D.E., Dorward, D.W., Whitmire, W.M., and Caldwell, H.D. (2010) Biological characterization of *Chlamydia trachomatis* plasticity zone MACPF domain

family protein CT153. *Infect Immun* **78**: 2691–2699.

Thalmann, J., Janik, K., May, M., Sommer, K., Ebeling, J., Hofmann, F., *et al.* (2010) Actin re-organization induced by *Chlamydia trachomatis* serovar D - evidence for a critical role of the effector protein CT166 targeting rac. *PLoS One* **5**: 1–14.

Thompson, C.C., Griffiths, C., Nicod, S.S., Lowden, N.M., Wigneshweraraj, S., Fisher, D.J., and McClure, M.O. (2015) The Rsb Phosphoregulatory Network Controls Availability of the Primary Sigma Factor in *Chlamydia trachomatis* and Influences the Kinetics of Growth and Development. *PLOS Pathog* **11**: e1005125.

Thomson, N.R., Holden, M.T.G., Carder, C., Lennard, N., Lockey, S.J., Marsh, P., *et al.* (2008) *Chlamydia trachomatis*: genome sequence analysis of lymphogranuloma venereum isolates. *Genome Res* **18**: 161–171.

Traven, A., and Naderer, T. (2014) Microbial Egress: A Hitchhiker's Guide to Freedom. *PLoS Pathog* **10**: 7–10.

Unemo, M., Seth-Smith, H.M.B., Cutcliffe, L.T., Skilton, R.J., Barlow, D., Goulding, D., *et al.* (2010) The Swedish new variant of *Chlamydia trachomatis*: Genome sequence, morphology, cell tropism and phenotypic characterization. *Microbiology* **156**: 1394–1404.

Valdivia, R.H. (2008) *Chlamydia* effector proteins and new insights into chlamydial cellular microbiology. *Curr Opin Microbiol* **11**: 53–59.

Vandahl, B.B., Birkelund, S., Demol, H., Hoorelbeke, B., Christiansen, G., Vandekerckhove, J., and Gevaert, K. (2001) Proteome analysis of the *Chlamydia pneumoniae*. *Electrophoresis* 1204–1223.

Vergne, I., Chua, J., Singh, S.B., and Deretic, V. (2004) Cell biology of *Mycobacterium tuberculosis* phagosome. *Annu Rev Cell Dev Biol* **20**: 367–394.

Voigt, A., Schöfl, G., and Saluz, H.P. (2012) The *Chlamydia psittaci* genome: A comparative analysis of intracellular pathogens. *PLoS One* **7**.

Vorimore, F., Hsia, R., Huot-Creasy, H., Bastian, S., Deruyter, L., Passet, A., *et al.* (2013) Isolation of a New *Chlamydia* species from the Feral Sacred Ibis (*Threskiornis aethiopicus*): *Chlamydia ibidis*. *PLoS One* **8**: 1–11.

Wang, Y., Kahane, S., Cutcliffe, L.T., Skilton, R.J., Lambden, P.R., and Clarke, I.N. (2011) Development of a transformation system for *Chlamydia trachomatis*: Restoration of glycogen biosynthesis by acquisition of a plasmid shuttle vector. *PLoS Pathog* **7**.

Wang, Y., Sun, M., Bao, H., and White, A.P. (2013) T3_MM: A Markov Model Effectively Classifies Bacterial Type III Secretion Signals. *PLoS One* **8**: 1–12.

Weber, M.M., Bauler, L.D., Lam, J., and Hackstadt, T. (2015) Expression and localization of predicted inclusion membrane proteins in *Chlamydia trachomatis*. *Infect Immun* IAI.01075–15.

Weitzman, M.D., and Weitzman, J.B. (2014) What's the damage? The impact of pathogens on pathways that maintain host genome integrity. *Cell Host Microbe* **15**: 283–294.

WHO (2008) Global incidence and prevalence of selected curable sexually transmitted infections. .

WHO (2014) Alliance for the Global Elimination of Blinding Trachoma by the year 2020. .

Wickstrum, J., Sammons, L.R., Restivo, K.N., and Hefty, P.S. (2013) Conditional gene expression in *Chlamydia trachomatis* using the tet system. *PLoS One* **8**: e76743.

Wolf, K., Plano, G. V., and Fields, K.A. (2009) A protein secreted by the respiratory pathogen *Chlamydia pneumoniae* impairs IL-17 signalling via interaction with human Act1. *Cell Microbiol* **11**: 769–79.

Wright, H.R., Turner, A., and Taylor, H.R. (2007) Trachoma. *Lancet* **371**: 1945–54.

Wu, X., Lei, L., Gong, S., Chen, D., Flores, R., and Zhong, G. (2011) The chlamydial periplasmic stress response serine protease cHtrA is secreted into host cell cytosol. *BMC Microbiol* **11**: 87.

Wylie, J.L., Hatch, G.M., and McClarty, G. (1997) Host cell phospholipids are trafficked to and then modified by *Chlamydia trachomatis*. *J Bacteriol* **179**: 7233–7242.

Xu, L., and Luo, Z.Q. (2013) Cell biology of infection by *Legionella pneumophila*. *Microbes Infect* **15**: 157–167.

Yao, J., Cherian, P.T., Frank, M.W., and Rock, C.O. (2015) *Chlamydia trachomatis* Relies on Autonomous Phospholipid Synthesis for Membrane Biogenesis. *J Biol Chem* **290**.

Zhong, G. (2011) *Chlamydia trachomatis* secretion of proteases for manipulating host signaling pathways. *Front Microbiol* **2**: 14.

CHAPTER 2 - Identification of type III secretion signals in inclusion membrane proteins of *Chlamydia trachomatis*

This Chapter contains unpublished data and data published in da Cunha, M., Milho, C., Almeida, F., Pais, S.V., Borges, V., Maurício, R., Borrego, M. J., Gomes, J.P., Mota, L.J. *(2014) Identification of type III secretion substrates of* Chlamydia trachomatis *using* Yersinia enterocolitica *as a heterologous system. BMC Microbiol* ***14****: 40, and in Almeida, F., Borges, V., Ferreira, R., Borrego, M.J., Gomes, J.P., and Mota, L.J. (2012) Polymorphisms in Inc proteins and differential expression of* inc *genes among* Chlamydia trachomatis *strains correlate with invasiveness and tropism of lymphogranuloma venereum isolates. J Bacteriol* ***194****: 6574–6585.*

The author of this dissertation participated in all experiments described in this Chapter.

2.1 Abstract

Chlamydia trachomatis multiplies exclusively within a membrane-bound vacuole (known as inclusion), and uses a type III secretion system to manipulate host cells by injecting them with effector proteins. In this Chapter, we identified type III secretion (T3S) signals in *C. trachomatis* inclusion membrane (Inc) proteins, a group of >60 proteins characterized by a bi-lobed hydrophobic domain that might mediate their insertion into the inclusion membrane. We focused our work in 23 known Incs (previously described to localize at the inclusion membrane) and in 25 putative Incs (not yet described to localize at the inclusion membrane, but containing the characteristic bi-lobed hydrophobic domain). We analyzed T3S of hybrid proteins comprising the first 20 amino acids of Inc proteins and the mature form of β-lactamase TEM-1, using *Yersinia enterocolitica* bacteria as heterologous host. Overall, we identified 19 Inc proteins of *C. trachomatis* as novel T3S substrates, of which 14 were putative Inc proteins. Bioinformatics further indicated the presence of a T3S signal in the first 20 amino acids of these Inc proteins. Our data supports the notion that most Inc proteins are T3S substrates and confirms that definitive indication of a protein as being an Inc requires immunolocalization studies in *Chlamydia*-infected cells.

2.2 Introduction

During its developmental cycle, *Chlamydia trachomatis* reside and multiply intracellularly in a membranous compartment known as an inclusion. Throughout this cycle, *C. trachomatis* uses a type III secretion system (T3SS) to translocate several effector proteins both across the host cell plasma membrane and the inclusion membrane (Valdivia, 2008; Betts *et al.*, 2009). *C. trachomatis* type III secretion (T3S) effector proteins should play a central role in bacterial invasion (Lane *et al.*, 2008; Hower *et al.*, 2009; Jewett *et al.*, 2010; Chen *et al.*, 2014), host cytoskeleton reorganization (Chen *et al.*, 2006; Dumoux *et al.*, 2015; Mital *et al.*, 2015), exit of host cells (Lutter *et al.*, 2013), and in subversion of many others host cell processes (Hackstadt *et al.*, 1999; Scidmore and Hackstadt, 2001; Belland *et al.*, 2003; Kumar *et al.*, 2006; Rzomp *et al.*, 2006; Chellas-Géry *et al.*, 2007; Delevoye *et al.*, 2008; Mital *et al.*, 2010; Pennini *et al.*, 2010; Derré *et al.*, 2011; Agaisse and Derré, 2014; Mirrashidi *et al.*, 2015). There are, however, chlamydial effector proteins, such as CPAF/CT858 or Tsp/CT441, which are not T3S substrates (Betts *et al.*, 2009).

Given their likely central role during infection, considerable efforts have been placed at identifying chlamydial effector proteins. This is not a trivial task because the amino acid sequence of most effectors does not display significant similarity to proteins of known function. Furthermore, a T3S signal is not evident on the basis of the amino acid sequence, predicted secondary structure, or other special properties. Moreover, in spite of the recent development of systems for transformation of *Chlamydia* (Wang *et al.*, 2011; Gérard *et al.*, 2013), for a long time no methods have been available for genetic manipulation of *Chlamydia*.

To overcome these obstacles chlamydial effectors have been searched: i) by systematic phenotypic analyses of yeast *Saccharomyces cerevisiae*

expressing individual chlamydial proteins (Sisko *et al.*, 2006); ii) by using *Salmonella* (Ho and Starnbach, 2005), *Shigella* (Subtil *et al.*, 2005; Pennini *et al.*, 2010; Dehoux *et al.*, 2011; Muschiol *et al.*, 2011; Furtado *et al.*, 2013), or *Yersinia* (Fields and Hackstadt, 2000; Clifton *et al.*, 2004; Chellas-Géry *et al.*, 2007; Hower *et al.*, 2009; Hovis *et al.*, 2013; Pais *et al.*, 2013) as genetically tractable heterologous host bacteria carrying well characterized T3SSs; or iii) by complex computational predictions of T3S signals (Arnold *et al.*, 2009; Löwer and Schneider, 2009; Samudrala *et al.*, 2009; Hovis *et al.*, 2013; Wang *et al.*, 2013). The subsequent use of specific antibodies enabled to detect translocation into host cells of some of the *C. trachomatis* proteins singled out in these searches, including a family of >60 proteins containing a hydrophobic motif thought to mediate their insertion into the inclusion membrane (Inc proteins) (Li *et al.*, 2008; Dehoux *et al.*, 2011). However, this distinct localization was only described for half of them (Li *et al.*, 2008; Mital *et al.*, 2010).

The majority of Inc proteins lacks predicted signal sequences required for secretion by other transport mechanisms, and therefore they are good candidates for being T3S effectors. Indeed, they appear to be transported into the inclusion membrane by a T3SS (Fields *et al*., 2003; Subtil *et al*., 2005; Dehoux *et al*., 2011). However, not all Incs were tested for the presence of a T3S signal. Aiming to test if it is possible to define a chlamydial protein as an Inc based on the presence of the typical hydrophobic motif and of a T3S signal, in this Chapter, we used *Y. enterocolitica* as a heterologous bacterium to identify T3S signals in putative Inc proteins of *C. trachomatis* (having a typical hydrophobic motif but not yet described to localize at the inclusion membrane) by comparison to known Inc proteins (described to localize at the inclusion membrane).

2.3 Materials and Methods

2.3.1 Bacterial strains and growth conditions

We used *Escherichia coli* TOP10 (Life Technologies) as host to help in the construction of plasmids and in the purification of the plasmid DNA. *Y. enterocolitica* ΔHOPEMT strain [MRS40 pYML421 (*yopH*$_{\Delta 1\text{-}352}$, *yopO*$_{\Delta 65\text{-}558}$ *yopP*$_{23}$ *yopE*$_{21}$ *yopM*$_{23}$ *yopT*$_{135}$)], deficient for the T3S effectors YopH, -O, -P, -E, -M, -T (Iriarte and Cornelis, 1998), but T3S-proficient, was used for T3S assays. To construct a T3S-deficient derivative of ΔHOPEMT, we deleted in this strain the complete coding sequence (codons 1 to 354) of the *yscU* gene, which encodes an essential component of the *Y. enterocolitica* T3SS (Sorg *et al.*, 2007). This was done by allelic exchange with the mutator plasmid pLY16. The *yscU*$_{\Delta 1\text{-}354}$ mutation had been previously shown to be non-polar (Sorg *et al.*, 2007). The resulting *Y. enterocolitica* ΔHOPEMT ΔYscU strain (MRS40 pFA1001) was also used in the T3S assays. *E. coli* and *Y. enterocolitica* were routinely grown at 37ºC and 28ºC, respectively, in LB medium with the appropriate antibiotics and supplements. Plasmids were routinely introduced into *E. coli* and *Y. enterocolitica* by electroporation.

2.3.2 Construction of plasmids

Plasmids were constructed and purified using proofreading Phusion DNA polymerase, restriction enzymes, T4 DNA Ligase, and DreamTaq DNA polymerase (all from Life Technologies), NucleoSpin gel and PCR cleanup kit (Macherey-Nagel) and GeneElute plasmid miniprep kit (Sigma). The accuracy of the nucleotide sequence of the inserts in the constructed plasmids was verified by DNA sequencing (Stab Vida). To analyze T3S signals we constructed plasmids harboring hybrid genes encoding the first 10, 15, 20, or 40 amino acids of each protein (*C.*

trachomatis Inc proteins, SycT and YopE) and the mature form of TEM-1 β-lactamase (TEM-1) (Charpentier and Oswald, 2004). These hybrids were constructed using as vector pLJM3, a low-copy plasmid which enables expression of the cloned genes driven by the promoter of the *Y. enterocolitica yopE* gene (Marenne *et al.*, 2003). This was done either by overlapping PCR between the first 60 nucleotides of each *inc* and the TEM-1-encoding gene or by using primers containing a restriction site followed by 60 nucleotides encoding the first 20 amino acids of each protein in frame with a sequence complementary to the 3' extremity of the transcribed strand of the TEM-1-encoding gene. Primers are listed in Table A.1 and plasmids are listed in Table A.2 (Annexes). For primer design, the DNA sequence of each *inc* gene in *C. trachomatis* strain L2/434 was used, except for *ct036*, *ct115* (*incD*), and *ct119* (*incA*), in which the sequence from strain D/UW3 was used (Table A.3 in Annexes). The gene encoding for TEM-1 was amplified by PCR from pCX340 (Charpentier and Oswald, 2004) and the DNA sequence of *sycT* was obtained from *Y. enterocolitica* pYVe227 (accession number AF102990).

2.3.3 Y. enterocolitica *T3S assays*

The T3S assays were performed as previously described (Sorg *et al.*, 2007). Briefly, *Y. enterocolitica* strains harboring the plasmids encoding TEM-1 hybrid proteins were diluted from overnight cultures to an OD_{600} of 0.1 in BHI medium supplemented with 20 mM sodium oxalate, 20 mM $MgCl_2$ and 0.4% (w/v) glucose, and grown for 2 h at 28ºC with 150 rpm shaking. The bacterial cultures were then quickly shifted to 37ºC and incubated for an additional 4 h with 150 rpm shaking. After incubation, OD_{600} of the cultures was measured and culture supernatants and bacterial pellets were separated by a centrifugation of 1 min, at room temperature (RT) and 17000 *g*. Proteins in the supernatant were

precipitated with TCA 10% (w/v). The TCA protein precipitate corresponding to 1 ml of the culture was resuspended in OD_{600} x 10 µl of SDS-PAGE Laemmli buffer [Tris-HCl 50 mM, pH 6.8, SDS 2.0% (w/v), glycerol 10% (v/v), β-mercaptoethanol 0.1 M, bromophenol blue 0.1% (w/v)], while the bacterial pellet, corresponding to 1.5 ml of the culture, was resuspended in OD_{600} x 100 µl of SDS-PAGE Laemmli Buffer. Both fractions were denatured at 100ºC for 10 min.

2.3.4 Immunoblotting

Supernatant and pellet fractions were resolved in 12% SDS-PAGE, followed both by Coomassie staining and by immunobloting. Proteins in the gel were transferred onto PVDF membranes (Bio-Rad) in a Trans-Blot® SD Semi-Dry System (Bio-Rad), blocked in 5% (w/v) dried skimmed milk diluted in PBS containing 0.1% (v/v) Tween-20. The membranes were probed with mouse anti-TEM1 antibody (QED Bioscience; used at 1/500) and rabbit anti-SycO antibody [(Letzelter *et al.*, 2006); used at 1/1000], horseradish peroxidase-conjugated secondary antibodies (GE Healthcare; used at 1:10000), and detected using the Western Lightning Plus-ECL kit (Perkin Elmer) in a Chemidoc XRS+ system (Bio-Rad). The amount of protein in the culture supernatant and bacterial pellet fractions was determined from the immunoblot images using Image Lab (Bio-Rad). The percentage of secretion was calculated as the ratio between the amounts of secreted protein (in the culture supernatant fraction) relative to the total amount of protein (in the culture supernatant and in the bacterial pellet fractions). The results from the quantifications are the average ± SEM from at least three independent experiments.

2.4 Results and discussion

2.4.1 TEM-1 β-lactamase can be used as a reporter protein to analyze T3S by Y. enterocolitica

The T3S signal is normally located within the first 20-30 amino acids of the substrate (Sory *et al.*, 1995; Samudrala *et al.*, 2009). Therefore, we analyzed secretion of hybrid proteins comprising the first amino acids of each *C. trachomatis* Inc protein and a reporter protein. We tested if the mature form of TEM-1 β-lactamase, which had been already described as a reporter protein for assays of T3S translocation into host cells (Charpentier and Oswald, 2004), could be used in the identification of T3S signals using *Y. enterocolitica* as a heterologous host. For this, we incubated in T3S-inducing conditions T3S-proficient *Y. enterocolitica* ΔHOPEMT or T3S-deficient *Y. enterocolitica* ΔHOPEMT ΔYscU strains harboring plasmids encoding an hybrid protein comprising the first 15 amino acids of the *Yersinia* effector YopE and TEM-1 ($YopE_{15}$-TEM-1) or TEM-1 alone. The first 15 amino acids of YopE have been previously shown to drive secretion of a Cya hybrid (Sory *et al.*, 1995). The different bacterial cultures were then fractionated into culture supernatants and bacterial pellets, which were subsequently analyzed by immunoblotting. This showed that the first 15 amino acids of YopE drove secretion of the TEM-1 hybrid, while mature TEM-1 alone was not secreted (Figure 2.1). Secretion of $YopE_{15}$-TEM-1 was T3S-dependent, as this protein was not detected in the supernatant fraction of the *Y. enterocolitica* ΔHOPEMT ΔYscU strain (Figure 2.1). This indicated that TEM-1 is a suitable reporter protein for the identification of T3S signals using *Y. enterocolitica* as a heterologous bacterium.

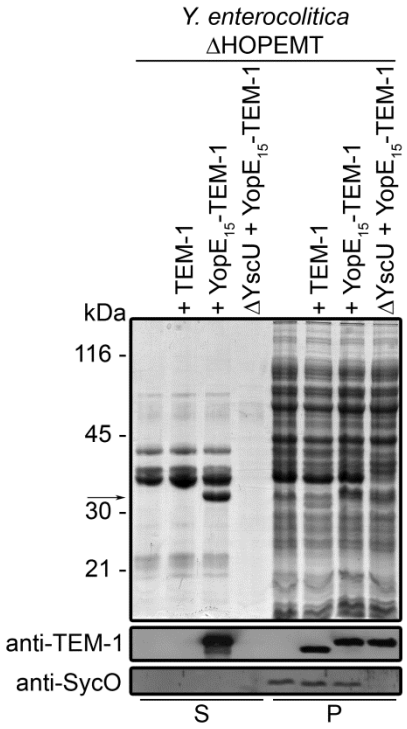


Figure 2.1 – Validation of mature TEM-1 β-lactamase as a reporter protein to identify T3S signals using *Y. enterocolitica* as a heterologous bacterium. The indicated *Y. enterocolitica* strains were incubated in T3S-inducing conditions (see Materials and Methods). Proteins in the supernatant (S) and in the bacterial pellet (P) were collected and analyzed by Coomassie staining (~$5x10^8$ and ~$1x10^8$ bacteria per lane for supernatant and pellet, respectively) and by immunoblotting (~$2x10^8$ and ~$2x10^7$ bacteria per lane for supernatant and pellet, respectively) with anti-TEM-1 and anti-SycO antibodies. SycO is a cytosolic T3S chaperone and was used as a control for possible contamination of the supernatant fractions with the bacterial pellet; SycO was not detected in the ΔHOPEMT ΔYscU background because its expression requires a functional T3SS (Stainier *et al.*, 1997).

2.4.2 The first 20 amino acids of C. trachomatis *T3S substrates are sufficient to drive efficient secretion of TEM-1 hybrid proteins by* Y. enterocolitica

Before analyzing T3S signals in Inc proteins, we sought to ascertain the optimal amino acid length of the chlamydial T3S signal that drives secretion of TEM-1 hybrid proteins in *Yersinia*. For this, we analyzed

secretion of proteins comprising the first 10, 20 and 40 amino acids of known *C. trachomatis* Incs that are T3S substrates (IncA or IncC) fused to TEM-1 ($IncA_{10}$-TEM-1, $IncA_{20}$-TEM-1, $IncA_{40}$-TEM-1, $IncC_{10}$-TEM-1, $IncC_{20}$-TEM-1, $IncC_{40}$-TEM-1) by *Y. enterocolitica* ΔHOPEMT or ΔHOPEMT ΔYscU (Figure 2.2). As negative controls we analyzed secretion by *Y. enterocolitica* ΔHOPEMT of TEM-1 alone and of a hybrid protein comprising the first 20 amino acids of the *Yersinia* T3S chaperone SycT and TEM-1 ($SycT_{20}$-TEM-1). As positive control, we analyzed secretion by ΔHOPEMT of $YopE_{15}$-TEM-1. Bacteria expressing these proteins were incubated in T3S-inducing conditions, and processed as described above (Figure 2.2). As observed before, mature TEM-1 alone was not secreted and, as expected, the $SycT_{20}$-TEM-1 fusion showed a basal percentage of secretion of 3.0 (SEM, 0.3). Based on this, to decide if a TEM-1 hybrid protein was secreted or not we set the threshold of percentage of secretion to 5.0 (Figure 2.2). We observed that the six Inc-TEM-1 hybrid proteins were type III secreted (Figure 2.2A and B). However, $IncA_{10}$-TEM-1 and $IncC_{10}$-TEM-1 were secreted less efficiently than $YopE_{15}$-TEM-1, while $IncA_{20}$-TEM-1, $IncA_{40}$-TEM-1, $IncC_{20}$-TEM-1 and $IncC_{40}$-TEM-1 were secreted at levels comparable to $YopE_{15}$-TEM-1 (Figure 2.2A). Overall, these experiments indicated that the first 20 amino acids of *C. trachomatis* T3S substrates are sufficient to drive secretion of TEM-1 hybrid proteins by *Y. enterocolitica* ΔHOPEMT as efficiently as the first 15 amino acids of the *Yersinia* effector YopE.

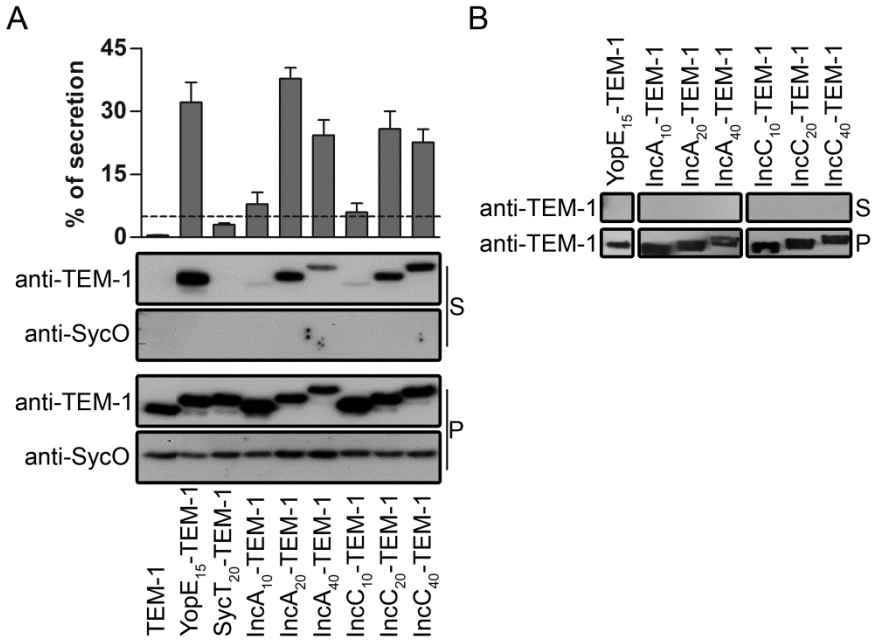


Figure 2.2 – The first 20 amino acids of known *C. trachomatis* T3S substrates (IncA or IncC) are sufficient to efficiently drive T3S of TEM-1 hybrid proteins by *Y. enterocolitica*. *Y. enterocolitica* T3S-proficient (ΔHOPEMT) (A) and T3S-deficient (ΔHOPEMT ΔYscU) (B) were used to analyze secretion of hybrid proteins comprising the first 10, 20, or 40 amino acids of *C. trachomatis* IncA or IncC, or the first 15 or 20 amino acids of *Y. enterocolitica* YopE or SycT, respectively, fused to the mature form of TEM-1 β-lactamase (TEM-1). Immunoblots show the result of T3S assays in which proteins in culture supernatants (S, secreted proteins) and in bacterial pellets (P, nonsecreted proteins) from ~$5x10^7$ bacteria were loaded per lane. $YopE_{15}$-TEM-1 was used as positive control, and TEM-1 and $SycT_{20}$-TEM-1 were used as negative controls for the T3S assays. The percentage (%) of secretion of each TEM-1 hybrid was calculated by densitometry, as the ratio between the amount of secreted and total protein. The threshold to decide whether a protein was secreted was set to 5.0% (dashed line), based on the % of secretion of $SycT_{20}$-TEM-1. Data are the mean ± SEM from at least 3 independent experiments.

2.4.3 The majority of Inc proteins of C. trachomatis *have a T3S signal recognized by the* Y. enterocolitica *T3SS*

We focused the identification of T3S signals in a group of 48 *C. trachomatis* known or predicted Inclusion membrane proteins studied by Li *et al*. (Table 2.1) (Li *et al.*, 2008). Only 23 of those 48 predicted Inc proteins have been detected at the inclusion membrane by

immunofluorescence microscopy using specific antibodies [referred to as known Inc proteins (Li *et al.*, 2008; Mital *et al.*, 2010; Dehoux *et al.*, 2011). The remaining 25 Inc proteins studied by Li and colleagues still lack conclusive experimental evidence about their localization in infected cells (referred to as putative Inc proteins) (Li *et al.*, 2008; Lutter *et al.*, 2012). We did not consider other chlamydial proteins previously described to be localized at the inclusion membrane (Cap1 and CopN), as they do not possess the characteristic bi-lobed hydrophobic domain of an Inc protein (Fields and Hackstadt, 2000; Fling *et al.*, 2001).

Table 2.1 – List of known and putative Inc proteins from *C. trachomatis* analyzed in this Chapter.

Inc protein[a]	T3S signal	Reference
Known Incs[b]		
CT101	No	This work
CT115/IncD	Yes	(Subtil *et al.*, 2005); this work
CT116/IncE	Yes	(Subtil *et al.*, 2005)
CT117/IncF	Yes	This work
CT118/IncG	Yes	(Subtil *et al.*, 2005)
CT119/IncA	Yes	(Subtil *et al.*, 2005); this work
CT147	ND[e]	-
CT222	Yes	This work
CT223	Yes	(Subtil *et al.*, 2005)
CT225	No	This work
CT226	Yes	(Dehoux *et al.*, 2011)
CT228	Yes	(Dehoux *et al.*, 2011); this work
CT229	Yes	(Subtil *et al.*, 2005)
CT232/IncB	Yes	This work
CT233/IncC	Yes	(Fields *et al.*, 2003; Subtil *et al.*, 2005); this work
CT249	Yes	(Dehoux *et al.*, 2011); this work
CT288	Yes	(Subtil *et al.*, 2005)
CT358	Yes	(Dehoux *et al.*, 2011); this work
CT440	Yes	(Dehoux *et al.*, 2011); this work
CT442	Yes	(Subtil *et al.*, 2005)
CT618	Yes	This work
CT813	Yes	This work
CT850	Yes/No[d]	(Dehoux *et al.*, 2011); this work
Putative Incs[c]		
CT005	Yes	This work

Inc protein[a]	T3S signal	Reference
CT006	ND[e]	-
CT036	Yes	This work
CT058	Yes	(Dehoux *et al.*, 2011); this work
CT134	No	This work
CT135	Yes	This work
CT164	No	This work
CT179	No	This work
CT192	No/Yes[d]	(Dehoux *et al.*, 2011); this work
CT195	Yes	(Dehoux *et al.*, 2011); this work
CT196	Yes	This work
CT214	Yes	This work
CT224	Yes	This work
CT227	Yes	This work
CT300	Yes	This work
CT345	Yes	This work
CT357	Yes	This work
CT365	Yes	This work
CT383	Yes	(Dehoux *et al.*, 2011); this work
CT449	Yes	This work
CT483	Yes	This work
CT484	No	(Dehoux *et al.*, 2011); this work
CT565	No	(Dehoux *et al.*, 2011); this work
CT728	No	This work
CT789	Yes	This work

[a]Proteins containing a bi-lobed hydrophobic motif that were analyzed by Li *et al* (Li *et al.*, 2008), which we selected to study in this work. We did not consider proteins (Cap1 and CopN) which localize at the inclusion membrane but which do not possess the bi-lobed hydrophobic domain (Fields and Hackstadt, 2000; Fling *et al.*, 2001). More recent bioinformatics-based analyses identified additional putative Inc proteins (CT018, CT079, CT081, CT244, CT324, CT326, CT556, CT578, CT616, CT618, CT642, CT645, CT788, CT789, CT814.1, CT819, CT837, CT846, and CT873) in *C. trachomatis* (Fields *et al.*, 2003; Subtil *et al.*, 2005; Dehoux *et al.*, 2011), but these proteins were not analyzed in this Chapter.

[b]Known Inc proteins contain a bi-lobed hydrophobic motif and have been localized at the inclusion membrane by immunofluorescence microscopy using specific antibodies (Li *et al.*, 2008; Mital *et al.*, 2010; Dehoux *et al.*, 2011).

[c]Putative Inc proteins contain a bi-lobed hydrophobic motif but have not yet been localized at the inclusion membrane.

[d]Conflicting data between our observations and previous analyses obtained by using *Shigella flexneri* as a heterologous bacterial host.

[e]ND, not determined.

We analyzed secretion of hybrid proteins comprising the first 20 amino acids of each putative or known Inc and TEM-1 by ΔHOPEMT or ΔHOPEMT ΔYscU *Y. enterocolitica* (Figure 2.3). The different *Y. enterocolitica* strains were incubated in T3S-inducing conditions, followed by fractionation of the bacterial cultures into culture supernatants and bacterial pellets and subsequent immunoblotting analyses of the proteins in the two fractions. In total, we have analyzed T3S signals in 24 putative Incs (18 have not been previously analyzed for T3S) and in 15 known Incs (7 have not been previously analyzed for T3S) (Figure 2.3 and Table 2.1). The expression levels of the TEM-1 hybrid of putative Inc CT006 were extremely low, which hampered the analysis of a T3S signal in this protein. These experiments led to the identification of a T3S signal in 18 putative Incs and in 12 known Incs (Figure 2.3 and Table 2.1). This revealed 5 known and 14 putative Incs as novel *C. trachomatis* T3S substrates (Figure 2.3 and Table 2.1). However, we did not detect a clear T3S signal in three known Incs (CT101, CT225, and CT850; Figure 2.3 and Table 2.1). Their T3S signal could extend beyond the first 20 amino acids or might not be recognized by the *Y. enterocolitica* T3S system; alternatively, they might be transported into the inclusion membrane by a distinct mechanism. Regardless of the exact explanation, this implies that the lack of a detectable T3S signal could not be taken as a definitive indication that a putative Inc does not localize at the inclusion membrane. Furthermore, the percentage of putative and known Incs analyzed whose first 20 amino acids could be recognized by the *Y. enterocolitica* T3S machinery was nearly identical [75% (18 out of 24) and 80% (12 out of 15), respectively]. Overall, these analyses indicated that most of the putative Incs analyzed have a T3S signal recognized by *Y. enterocolitica* and therefore are likely T3S substrates, and support the notion that the vast

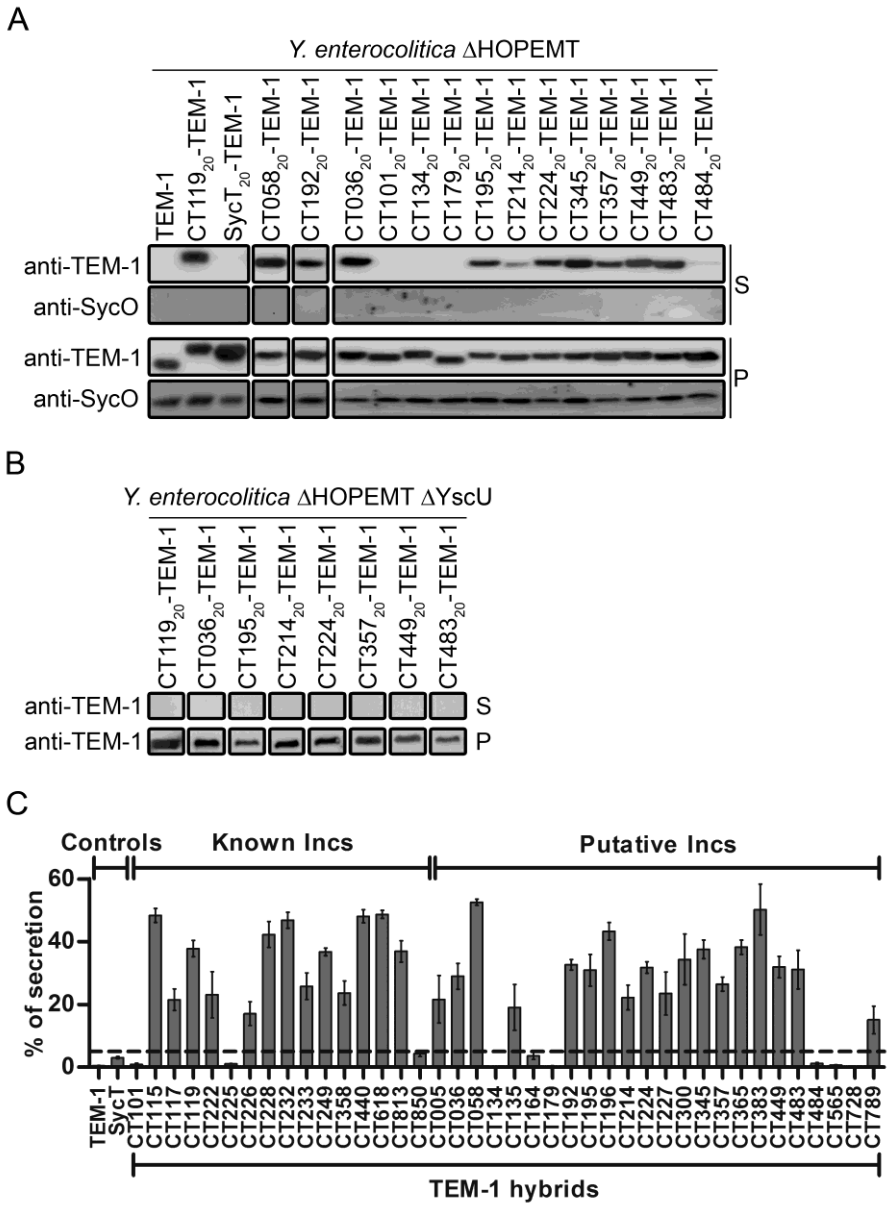


Figure 2.3 – Type III secretion (T3S) signals in *C. trachomatis* Inc proteins. *Y. enterocolitica* T3S-proficient (ΔHOPEMT) and T3S-deficient (ΔHOPEMT ΔYscU) bacteria were used to analyze secretion of hybrid proteins comprising the first 20 amino acids of Incs, or of *Y. enterocolitica* SycT, fused to the mature form of TEM-1 β-lactamase (TEM-1). (A) and (B) Immunoblots show the result of representative assays in which proteins in culture supernatants (S - secreted proteins), and in bacterial pellets (P - nonsecreted proteins) from ~5x10^7 bacteria were loaded per lane. SycT and SycO are strictly cytosolic *Yersinia* T3S chaperones (Iriarte and Cornelis, 1998; Letzelter *et al.*, 2006). SycT$_{20}$-TEM-1 was a negative control for the T3S assays. Immunodetection of

SycO ensured that presence of TEM-1 hybrid proteins in culture supernatants was not a result of bacterial lysis or contamination. (C) The percentage (%) of secretion of each TEM-1 hybrid was calculated by densitometry, as the ratio between the amount of secreted and total protein. The threshold to decide whether a protein was secreted was set to 5% (dashed line), based on the % of secretion of $SycT_{20}$-TEM-1. Data are the mean ± SEM from at least 3 independent experiments. Please note that in this work we did not analyze all described putative Incs (see legend of Table 2.1) and that T3S signals were not analyzed for all known Incs.

majority of Inc proteins are T3S effectors. Our work also shows that the analysis of T3S signals in proteins possessing a bi-lobed hydrophobic domains suggesting that they are Incs cannot be used to deduce whether they localize at the inclusion membrane. Instead, specific immunofluorescence analyses of *Chlamydia*-infected cells are required for such identification.

We analyzed the presence of a T3S signal in 14 Incs that have been previously analyzed by others (Table 2.1) (Fields *et al.*, 2003; Subtil *et al.*, 2005; Dehoux *et al.*, 2011). Our results were in agreement with the published data except for two proteins. CT192 was secreted by *Yersinia* in our study as a TEM-1 hybrid but not as a Cya hybrid by *Shigella* (Dehoux *et al.*, 2011). CT850 was not secreted by *Yersinia* in our study as a TEM-1 hybrid but was secreted as a Cya hybrid by *Shigella* (Dehoux *et al.*, 2011). These results might be explained by the different length of the amino acid sequences used in the fusions of TEM-1 and Cya, and also differences in the recognition of the *C. trachomatis* T3S signal by *Yersinia* and *Shigella*.

2.4.4 Prediction of T3S signals in C. trachomatis *Inc proteins*

In the last 6 years, several algorithms have been developed aimed at identifying T3S substrates *in silico* (Arnold *et al.*, 2009; Löwer and Schneider, 2009; Samudrala *et al.*, 2009; Hovis *et al.*, 2013; Wang *et al.*,

2013). Arnold and colleagues developed a machine learning tool [Effective; (Arnold *et al.*, 2009)] that searches for T3S signals based on the protein secondary structure, and frequency and physico-chemical properties of amino acids in the N-terminal region of known T3S substrates. Lower and colleagues developed a neural network [Modlab; (Löwer and Schneider, 2009) that scans the entire protein for common features shared by a dataset of known T3S effectors. Samudrala and colleagues created a machine learning tool [Sieve; (Samudrala *et al.*, 2009)] to identify T3S signals based on evolutionary relationships, amino acid composition, and G+C content of the first 30 amino acid residues of a protein. Hovis and colleagues modified this last tool to create another machine learning tool [cSieve; (Hovis et al., 2013)] that includes *Chlamydia* genus specific properties to better predict *C. trachomatis* T3S effectors. Wang and colleagues developed a Hidden-Markov model [T3_MM; (Wang *et al.*, 2013)] that analyzes not only the amino acid composition, but also the preferable position of each amino acid in relation to the others in the N-terminal 100 residues.

In this Chapter, we analyzed in *Yersinia* the secretion of 39 different TEM-1 hybrid proteins directed by the first 20 amino acids of *C. trachomatis* Inc proteins. We obtained the prediction of secretion for the TEM-1 hybrid proteins using the 3 prediction tools available online (Effective, Modlab, and T3_MM), and retrieved the list of Inc proteins from the chlamydial T3S substrates predicted by cSieve (Hovis et al., 2013) (Table 2.2). We also compared our data and the predictions to previous T3S assays using *Shigella* as heterologous bacteria (Subtil *et al.*, 2005; Dehoux *et al.*, 2011).

Overall, out of the 39 TEM-1 hybrids analyzed in this Chapter, Effective predicted 26 (66.7%), Modlab 23 (59.0%), and T3_MM 10 (25.6%) as being T3S substrates. In relation to our data, out of the 30 TEM-1 hybrid proteins in which a T3S signal was recognized by *Yersinia*, Effective

predicted 24 (80.0%), Modlab 22 (73.3%), and T3_MM 8 (26.7%) as possessing a T3S signal. The cSieve tool predicted 22 of the 39 (56.4%) Incs analyzed in this Chapter as TEM-1 hybrids as being T3S substrates, and 19 out of 30 Incs (63.3%) whose first 20 amino acids drive T3S of TEM-1 hybrids by *Yersinia* were among the chlamydial T3S substrates predicted by cSieve.

Specifically, and in agreement with our work, 6 out of 30 Incs (20.0%; CT228, CT358, CT813, CT345, CT357, and CT383) whose first 20 amino acids drove T3S of TEM-1 hybrids by *Yersinia* were predicted to have a T3S signal by all methods (Effective, Modlab, T3_MM, and cSieve). In addition, ten (33.3%) and eight (26.7%) of the TEM-1 hybrids secreted in *Yersinia* were predicted to have a T3S signal by 3 or 2, respectively, prediction tools simultaneously. Moreover, we found 5 out of 9 (55.6%) TEM-1 hybrids that were not secreted by *Yersinia* and for which the predictions also indicated that they do not have a T3S signal (the first 20 amino acids of CT225, CT850, CT134, CT179, and CT484). In these proteins, the T3S signal could extend beyond the 20 amino acids and might be truncated due to the fusion to the TEM-1 protein.

Table 2.2 – Comparison of the T3S signals identified in this Chapter with *in silico* predictions and results from other studies.

Inc protein[a]	Effective[b]	Modlab[c]	T3_MM[d]	cSieve[e]	Dehoux[f]	*Y. enterocolitica*[g]
Known Incs[h]						
CT101	No	No	No	Yes	NT[j]	No
CT115/IncD	Yes	No	No	No	Yes	Yes
CT117/IncF	No	No	No	No	NT	Yes
CT119/IncA	Yes	Yes	Yes	No	Yes	Yes
CT222	Yes	Yes	No	Yes	NT	Yes
CT225	No	No	No	No	NT	No
CT228	Yes	Yes	Yes	Yes	Yes	Yes
CT232/IncB	Yes	No	No	Yes	NT	Yes
CT233/IncC	Yes	Yes	No	Yes	Yes	Yes
CT249	No	No	No	Yes	Yes	Yes
CT358	Yes	Yes	Yes	Yes	Yes	Yes

CT440	Yes	Yes	No	Yes	Yes	Yes
CT618	Yes	Yes	No	No	NT	Yes
CT813	Yes	Yes	Yes	Yes	NT	Yes
CT850	No	No	No	No	Yes	No
Putative Incs[i]						
CT005	Yes	Yes	Yes	No	NT	Yes
CT036	No	No	No	Yes	NT	Yes
CT058	Yes	Yes	No	Yes	Yes	Yes
CT134	No	No	No	No	NT	No
CT135	Yes	Yes	No	No	NT	Yes
CT164	Yes	Yes	Yes	Yes	NT	No
CT179	No	No	No	No	NT	No
CT192	Yes	Yes	No	No	No	Yes
CT195	No	No	No	No	Yes	Yes
CT196	Yes	Yes	No	Yes	NT	Yes
CT214	No	Yes	No	Yes	NT	Yes
CT224	Yes	Yes	No	No	NT	Yes
CT227	Yes	Yes	No	No	NT	Yes
CT300	Yes	Yes	No	Yes	NT	Yes
CT345	Yes	Yes	Yes	Yes	NT	Yes
CT357	Yes	Yes	Yes	Yes	NT	Yes
CT365	Yes	Yes	No	Yes	NT	Yes
CT383	Yes	Yes	Yes	Yes	Yes	Yes
CT449	Yes	No	No	Yes	NT	Yes
CT483	Yes	Yes	No	Yes	NT	Yes
CT484	No	No	No	No	No	No
CT565	Yes	No	Yes	No	No	No
CT728	No	No	No	Yes	NT	No
CT789	No	No	No	No	NT	Yes

[a]Inc proteins analyzed as TEM-1 hybrids in this Chapter.

[b]Inc proteins predicted to be secreted as TEM-1 hybrids by Effective (Arnold *et al.*, 2009) using a cutoff of 0.6.

[c]Inc proteins predicted to be secreted as TEM-1 hybrids by Modlab (Löwer and Schneider, 2009) using a cutoff of 0.4.

[d]Inc proteins predicted to be secreted as TEM-1 hybrids by T3_MM (Wang *et al.*, 2013) scoring a positive value.

[e]Inc proteins predicted to have a T3S signal by cSieve (Hovis *et al.*, 2013).

[f]Cya hybrids that were secreted by *Shigella* (Dehoux *et al.*, 2011).

[g]TEM-1 hybrid proteins that were secreted by *Y. enterocolitica* in this Chapter.

[h]Inc proteins previously shown lo localize at the inclusion membrane.

[i]Inc proteins not yet described to localize at the inclusion membrane.

[j]NT not tested.

On the other hand, we found 3 out of 30 Incs (10.0%; CT117, CT195, and CT789) whose first 20 amino acids drove T3S of TEM-1 hybrids by *Yersinia* but for which all predictions failed to detect a T3S signal. Furthermore, among the 9 Incs whose first 20 amino acids did not drive T3S of TEM-1 hybrids by *Yersinia*, one (11.1%, CT164) was predicted to have a T3S signal by all 4 methods, and two (22.2%; CT101 and CT728) and one (11.1%; CT565) were predicted not to have a T3S signal by 3 or 2, respectively, prediction tools.

In summary, for 3 Incs (CT228, CT358, and CT383), data from *Yersinia*, *Shigella*, and the predictions indicate that these proteins are T3S substrates, and on the other hand, we found one Inc (CT484) that was not considered a T3S substrate by any of the analyses mentioned above. In general, Effective was the strongest prediction tool, generating a higher number of T3S substrates (80.0% of the 30 *Yersinia* secreted TEM-1 hybrids), although Modlab also gave similar results (~70.0%). Only eight of the 39 Incs (20.5%) were not predicted to have a T3S signal by any of the 4 prediction tools, indicating that these prediction tools have a good predictive potential but still there is room for improvement. The length of the N-terminal used in the fusions to the reporter protein also influences these predictions as in some cases the T3S signal could extend beyond the 20 amino acids considered as the prototypical length of the T3S signal.

2.4.5 Composition analysis of the first 20 amino acids of the N-terminal of Inc proteins of C. trachomatis

Although previous studies did not reveal any obvious feature defining a universal T3S signal, we analyzed our data for common features. The comparison between known T3S effectors and non-T3S effectors has given valuable insights about the nature of the T3S signal. Analyses of

the amino acid composition of the N-terminal regions of these proteins indicated several characteristics that were crucial to develop the prediction tools mentioned above (Arnold *et al.*, 2009; Löwer and Schneider, 2009; Samudrala *et al.*, 2009; Hovis *et al.*, 2013; Wang *et al.*, 2013). Among these characteristics, we focused on enrichment or depletion of specific amino acids or groups of amino acids. We calculated the total amino acid composition of the N-terminal 20 amino acids of the 30 Inc proteins that were secreted as TEM-1 hybrids in this Chapter (Figure 2.4, Figure 2.5 and Figure 2.6). Like reported before, we observed a predominant frequency of serine and proline (14.5% and 11.0%, respectively, of all the amino acid composition of the first 20 amino acids of the 30 secreted TEM-1 Inc hybrids), and to a less extent of threonine (7.8%). The observed reduced frequency of lysine and glutamate (2.3% and 2.7%, respectively) were also consistent with the previous data, however, the reduced frequency in leucine (6.7%) was not a feature in our set of proteins. We could also see a reduced frequency of charged residues (arginine, histidine, tyrosine, cysteine, and aspartate; the sum is 14.2%). Analyzing the position of the amino acids in the sequence we observed that hydrophobic (non-polar) residues were more scattered and alternated with hydrophilic (polar) residues, also in agreement with what was previously reported (Figure 2.4).

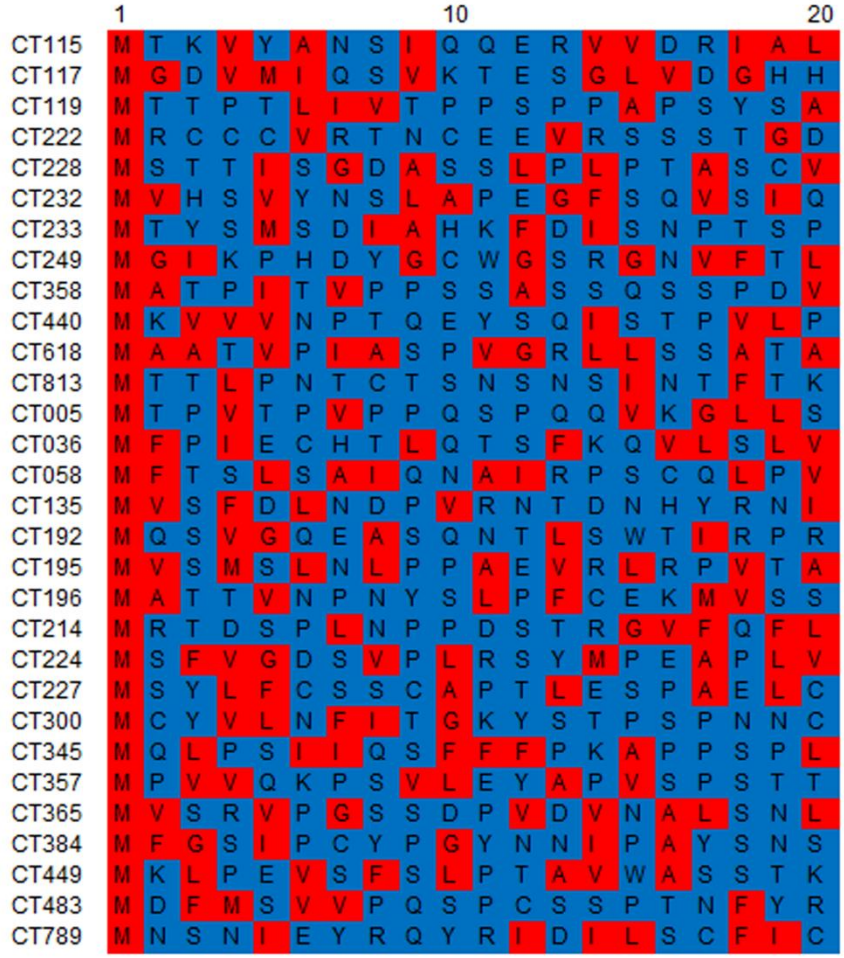


Figure 2.4 – Unaligned N-terminal sequences of the 30 Inc proteins that were secreted as TEM-1 hybrids in this Chapter. Hydrophilic or polar residues (C, D, E, H, K, N, P, Q, R, S, T, W and Y) are colored in blue and hydrophobic or non-polar residues (A, F, G, I, L, M and V) are colored in red.

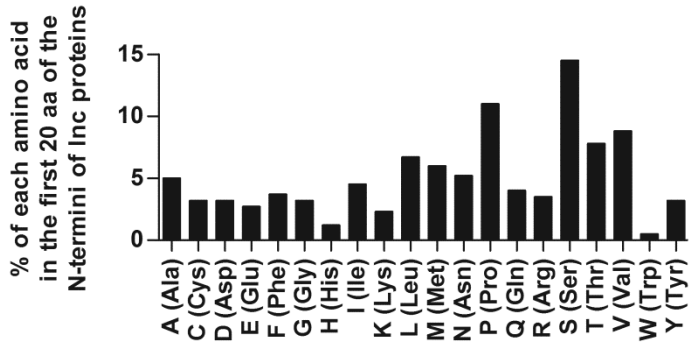


Figure 2.5 – Composition analysis of the 20 amino acids at the N-termini of Inc proteins. Total amino acid composition of the first 20 amino acids of the 30 Inc proteins that were secreted as TEM-1 hybrids in this Chapter.

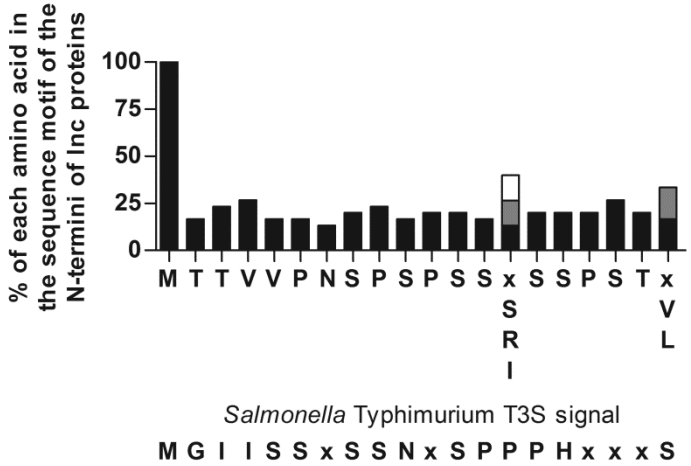


Figure 2.6 – Consensus-like sequence motif of the T3S signal present in Inc proteins. Sequence motif of the T3S signal of the 30 Inc proteins that were secreted as TEM-1 hybrids in this Chapter, showing the amino acids with the highest positive weight for each amino acid position, from 1 to 20. For comparison we show below the sequence motif for *S.* Typhimurium T3S effectors (Samudrala *et al.*, 2009), showing the amino acids with the highest positive weight for each amino acid position, from 1 to 20. Positions with an ‘‘x’’ have no representation in the sequence motif.

Samudrala and colleagues (Samudrala *et al.*, 2009) analyzed the N-terminal of effectors from *Salmonella enterica* serovar Typhimurium (*S.* Typhimurium) and *Pseudomonas syringae* and were able to derive a consensus-like sequence motif. We compared the sequence motif found for *Salmonella* effectors with the sequence motif found in this Chapter for the Inc proteins from *C. trachomatis* that were secreted as TEM-1 hybrids (Figure 2.4 and Figure 2.6). Besides the methionine in the first position, both consensus sequences had a serine in the 8^{th} and 12^{th} position. Although a conserved motif was not found between the two species, some similarities could be identified. Both sequence motifs are highly represented (only 2 and 4 positions for the sequence motif of *C. trachomatis* Inc proteins and *Salmonella* effectors, respectively, with no representation).

2.5 Conclusions

To understand if the presence or lack of a T3S signal together with a hydrophobic domain characteristic of Incs could be used to identify a protein as an Inc or not, we analyzed T3S signals in a pool of *C. trachomatis* known and putative Inc proteins. As we were unable to identify a T3S signal in a few known Inc proteins (CT101, CT225, and CT850), we conclude that definitive identification of a protein as an Inc requires their immunolocalization at the inclusion membrane in *Chlamydia*-infected cells. However, these analyses led to the identification of 19 novel T3S substrates, including 14 putative Inc proteins. Additionally, data from 4 T3S signal prediction tools available in the literature and analyses of the amino acid composition of the N-terminal of Inc proteins revealed common characteristics to known T3S effectors of other bacteria like *Salmonella*. Our data supports the notion that the majority of Inc proteins possess a T3S signal that enables the translocation into the host cell. It remains to be understood how after crossing the inclusion membrane and reaching the host cell cytosol the Inc proteins are inserted into the inclusion membrane. The characteristic hydrophobic domain likely plays a fundamental role in this process, but the mechanistic details are still unknown.

2.6 Acknowledgements

The work reported in this Chapter was supported by Fundação para a Ciência e a Tecnologia (FCT) through grants PEst-OE/EQB/LA0004/2011, PTDC/SAU-MII/099623/2008 and ERA-PTG/0005/2010 (in the frame of ERA-NET PathoGenoMics); by the European Commission through a Marie Curie European Re-integration Grant (PERG03-GA-2008-230954); and by a European Society for Clinical Microbiology and Infectious Diseases (ESCMID) research grant. Filipe Almeida was a recipient of a PhD fellowship from FCT (SFRH/BD/73545/2010).

2.7 References

Agaisse, H., and Derré, I. (2014) Expression of the effector protein IncD in *Chlamydia trachomatis* mediates recruitment of the lipid transfer protein CERT and the endoplasmic reticulum-resident protein VAPB to the inclusion membrane. *Infect Immun* **82**: 2037–2047.

Arnold, R., Brandmaier, S., Kleine, F., Tischler, P., Heinz, E., Behrens, S., *et al.* (2009) Sequence-Based Prediction of Type III Secreted Proteins. *PLoS Pathog* **5**.

Belland, R.J., Zhong, G., Dean, D., Hogan, D., Sturdevant, D.E., Sharma, J., *et al.* (2003) Genomic transcriptional profiling of the developmental cycle of *Chlamydia trachomatis*. *Proc Natl Acad Sci U S A* **100**: 8478–8483.

Betts, H.J., Wolf, K., and Fields, K.A. (2009) Effector protein modulation of host cells: examples in the *Chlamydia* spp. arsenal. *Curr Opin Microbiol* **12**: 81–87.

Charpentier, X., and Oswald, E. (2004) Identification of the secretion and translocation domain of the enteropathogenic and enterohemorrhagic *Escherichia coli* effector Cif, using TEM-1 beta-lactamase as a new fluorescence-based reporter. *J Bacteriol* **186**: 5486–5495.

Chellas-Géry, B., Linton, C.N., and Fields, K.A. (2007) Human GCIP interacts with CT847, a novel *Chlamydia trachomatis* type III secretion substrate, and is degraded in a tissue-culture infection model. *Cell Microbiol* **9**: 2417–2430.

Chen, C., Chen, D., Sharma, J., Cheng, W., Zhong, Y., Liu, K., *et al.* (2006) The hypothetical protein CT813 is localized in the *Chlamydia trachomatis* inclusion membrane and is immunogenic in women urogenitally infected with *C. trachomatis*. *Infect Immun* **74**: 4826–4840.

Chen, Y.-S., Bastidas, R.J., Saka, H.A., Carpenter, V.K., Richards, K.L., Plano, G. V., and Valdivia, R.H. (2014) The *Chlamydia trachomatis* Type III Secretion Chaperone Slc1 Engages Multiple Early Effectors, Including TepP, a Tyrosine-phosphorylated Protein Required for the Recruitment of CrkI-II to Nascent Inclusions and Innate Immune Signaling. *PLoS Pathog* **10**.

Clifton, D.R., Fields, K.A., Grieshaber, S.S., Dooley, C.A., Fischer, E.R., Mead, D.J., *et al.* (2004) A chlamydial type III translocated protein is tyrosine-phosphorylated at the site of entry and associated with recruitment of actin. *Proc Natl Acad Sci U S A* **101**: 10166–10171.

Dehoux, P., Flores, R., Dauga, C., Zhong, G., and Subtil, A. (2011) Multi-genome identification and characterization of *Chlamydiae*-specific type III secretion substrates: the Inc proteins. *BMC Genomics* **12**: 109.

Delevoye, C., Nilges, M., Dehoux, P., Paumet, F., Perrinet, S., Dautry-Varsat, A., and Subtil, A. (2008) SNARE protein mimicry by an intracellular bacterium. *PLoS Pathog* **4**.

Derré, I., Swiss, R., and Agaisse, H. (2011) The lipid transfer protein CERT interacts with the *Chlamydia* inclusion protein IncD and participates to ER-*Chlamydia* inclusion membrane contact sites. *PLoS Pathog* **7**.

Dumoux, M., Menny, A., Delacour, D., and Hayward, R.D. (2015) A *Chlamydia* effector recruits CEP170 to reprogram host microtubule organization. *J Cell Sci* **44**.

Fields, K.A., and Hackstadt, T. (2000) Evidence for the secretion of *Chlamydia trachomatis* CopN by a type III secretion mechanism. *Mol Microbiol* **38**: 1048–1060.

Fields, K.A., Mead, D.J., Dooley, C.A., and Hackstadt, T. (2003) *Chlamydia trachomatis*

type III secretion: Evidence for a functional apparatus during early-cycle development. *Mol Microbiol* **48**: 671–683.

Fling, S.P., Sutherland, R. a, Steele, L.N., Hess, B., D'Orazio, S.E., Maisonneuve, J., *et al.* (2001) CD8+ T cells recognize an inclusion membrane-associated protein from the vacuolar pathogen *Chlamydia trachomatis*. *Proc Natl Acad Sci U S A* **98**: 1160–1165.

Furtado, A.R., Essid, M., Perrinet, S., Balañá, M.E., Yoder, N., Dehoux, P., and Subtil, A. (2013) The chlamydial OTU domain-containing protein ChlaOTU is an early type III secretion effector targeting ubiquitin and NDP52. *Cell Microbiol* **15**: 2064–2079.

Gérard, H.C., Mishra, M.K., Mao, G., Wang, S., Hali, M., Whittum-Hudson, J. a., *et al.* (2013) Dendrimer-enabled DNA delivery and transformation of *Chlamydia pneumoniae*. *Nanomedicine Nanotechnology, Biol Med* **9**: 996–1008.

Hackstadt, T., Scidmore, M.A., Shaw, E.I., and Fischer, E.R. (1999) The *Chlamydia trachomatis* IncA protein is required for homotypic vesicle fusion. *Cell Microbiol* **1**: 119–130.

Ho, T.D., and Starnbach, M.N. (2005) The *Salmonella enterica* serovar Typhimurium-encoded type III secretion systems can translocate *Chlamydia trachomatis* proteins into the cytosol of host cells. *Infect Immun* **73**: 905–911.

Hovis, K.M., Mojica, S., Mcdermott, J.E., Pedersen, L., Simhi, C., Rank, R.G., *et al.* (2013) Genus-optimized strategy for the identification of chlamydial type III secretion substrates. *Pathog Dis* **69**: 213–222.

Hower, S., Wolf, K., and Fields, K.A. (2009) Evidence that CT694 is a novel *Chlamydia trachomatis* T3S substrate capable of functioning during invasion or early cycle development. *Mol Microbiol* **72**: 1423–1437.

Iriarte, M., and Cornelis, G.R. (1998) YopT , a new *Yersinia* Yop effector protein , affects the cytoskeleton of host cells. *Mol Biosyst* **29**: 915–929.

Jewett, T.J., Miller, N.J., Dooley, C.A., and Hackstadt, T. (2010) The conserved tarp actin binding domain is important for chlamydial invasion. *PLoS Pathog* **6**: 1–11.

Kumar, Y., Cocchiaro, J., and Valdivia, R.H. (2006) The Obligate Intracellular Pathogen *Chlamydia trachomatis* Targets Host Lipid Droplets. *Curr Biol* **16**: 1646–1651.

Lane, B.J., Mutchler, C., Khodor, S. Al, Grieshaber, S.S., and Carabeo, R.A. (2008) Chlamydial entry involves TARP binding of guanine nucleotide exchange factors. *PLoS Pathog* **4**.

Letzelter, M., Sorg, I., Meyer, S., Feldman, M.F., Kuhn, M., Callebaut, I., and Cornelis, G.R. (2006) The discovery of SycO highlights a new function for type III secretion effector chaperones. *EMBO J* **25**: 3223–3233.

Li, Z., Chen, C., Chen, D., Wu, Y., Zhong, Y., and Zhong, G. (2008) Characterization of fifty putative inclusion membrane proteins encoded in the *Chlamydia trachomatis* genome. *Infect Immun* **76**: 2746–2757.

Löwer, M., and Schneider, G. (2009) Prediction of type III secretion signals in genomes of gram-negative bacteria. *PLoS One* **4**: 1435–1439.

Lutter, E.I., Barger, A.C., Nair, V., and Hackstadt, T. (2013) *Chlamydia trachomatis* Inclusion Membrane Protein CT228 Recruits Elements of the Myosin Phosphatase Pathway to Regulate Release Mechanisms. *Cell Rep* **3**: 1921–1931.

Lutter, E.I., Martens, C., and Hackstadt, T. (2012) Evolution and conservation of predicted inclusion membrane proteins in *Chlamydiae*. *Comp Funct Genomics* **2012**.

Marenne, M.N., Journet, L., Mota, L.J., and Cornelis, G.R. (2003) Genetic analysis of the formation of the Ysc-Yop translocation pore in macrophages by *Yersinia enterocolitica*: Role of LcrV, YscF and YopN. *Microb Pathog* **35**: 243–258.

Mirrashidi, K.M., Elwell, C.A., Verschueren, E., Johnson, J.R., Frando, A., Dollen, J. Von, *et al.* (2015) Global Mapping of the Inc-Human Interactome Reveals that Retromer Restricts *Chlamydia* Infection. *Cell Host Microbe* **18**: 109–121.

Mital, J., Lutter, E.I., Barger, A.C., Dooley, C.A., and Hackstadt, T. (2015) *Chlamydia trachomatis* inclusion membrane protein CT850 interacts with the dynein light chain DYNLT1 (Tctex1). *Biochem Biophys Res Commun* **1**: 1–6.

Mital, J., Miller, N.J., Fischer, E.R., and Hackstadt, T. (2010) Specific chlamydial inclusion membrane proteins associate with active Src family kinases in microdomains that interact with the host microtubule network. *Cell Microbiol* **12**: 1235–1249.

Muschiol, S., Boncompain, G., Vromman, F., Dehoux, P., Normark, S., Henriques-Normark, B., and Subtil, A. (2011) Identification of a family of effectors secreted by the type III secretion system that are conserved in pathogenic *Chlamydiae*. *Infect Immun* **79**: 571–580.

Pais, S. V., Milho, C., Almeida, F., and Mota, L.J. (2013) Identification of Novel Type III Secretion Chaperone-Substrate Complexes of *Chlamydia trachomatis*. *PLoS One* **8**.

Pennini, M.E., Perrinet, S., Dautry-Varsat, A., and Subtil, A. (2010) Histone methylation by NUE, a novel nuclear effector of the intracellular pathogen *Chlamydia trachomatis*. *PLoS Pathog* **6**: 1–12.

Rzomp, K.A., Moorhead, A.R., and Scidmore, M.A. (2006) The GTPase Rab4 interacts with *Chlamydia trachomatis* inclusion membrane protein CT229. *Infect Immun* **74**: 5362–5373.

Samudrala, R., Heffron, F., and Mcdermott, J.E. (2009) Accurate Prediction of Secreted Substrates and Identification of a Conserved Putative Secretion Signal for Type III Secretion Systems. *PLoS Pathog* **5**.

Scidmore, M.A., and Hackstadt, T. (2001) Mammalian 14-3-3-beta associates with the *Chlamydia trachomatis* inclusion membrane via its interaction with IncG. *Mol Microbiol* **39**: 1638–1650.

Sisko, J.L., Spaeth, K.E., Kumar, Y., and Valdivia, R.H. (2006) Multifunctional analysis of *Chlamydia*-specific genes in a yeast expression system. *Mol Microbiol* **60**: 51–66.

Sorg, I., Wagner, S., Amstutz, M., Müller, S. a, Broz, P., Lussi, Y., *et al.* (2007) YscU recognizes translocators as export substrates of the *Yersinia* injectisome. *EMBO J* **26**: 3015–3024.

Sory, M., Boland, A., Lambermont, I., and Cornelis, G.R. (1995) Identification of the YopE and YopH domains required for secretion and internalization into the cytosol of macrophages , using the *cyaA* gene fusion approach. *Proc Natl Acad Sci U S A* **92**: 11998–12002.

Stainier, I., Iriarte, M., and Cornelis, G.R. (1997) YscM1 and YscM2, two *Yersinia enterocolitica* proteins causing downregulation of yop transcription. *Mol Microbiol* **26**: 833–843.

Subtil, A., Delevoye, C., Balañá, M.E., Tastevin, L., Perrinet, S., and Dautry-Varsat, A. (2005) A directed screen for chlamydial proteins secreted by a type III mechanism identifies a translocated protein and numerous other new candidates. *Mol Microbiol* **56**: 1636–1647.

Valdivia, R.H. (2008) *Chlamydia* effector proteins and new insights into chlamydial cellular microbiology. *Curr Opin Microbiol* **11**: 53–59.

Wang, Y., Kahane, S., Cutcliffe, L.T., Skilton, R.J., Lambden, P.R., and Clarke, I.N. (2011) Development of a transformation system for *Chlamydia trachomatis*: Restoration of glycogen biosynthesis by acquisition of a plasmid shuttle vector. *PLoS Pathog* **7**.

Wang, Y., Sun, M., Bao, H., and White, A.P. (2013) T3_MM: A Markov Model Effectively Classifies Bacterial Type III Secretion Signals. *PLoS One* **8**: 1–12.

CHAPTER 3 - Polymorphisms in Inc proteins and differential expression of *inc* genes among *Chlamydia trachomatis* strains correlate with invasiveness and tropism of lymphogranuloma venereum isolates

This Chapter contains data published in Almeida, F., Borges, V., Ferreira, R., Borrego, M.J., Gomes, J.P., and Mota, L.J. (2012) Polymorphisms in Inc proteins and differential expression of inc *genes among* Chlamydia trachomatis *strains correlate with invasiveness and tropism of lymphogranuloma venereum isolates. J Bacteriol* ***194****: 6574–6585.*

The author of this dissertation participated in all experiments described in this Chapter.

3.1 Abstract

Chlamydia trachomatis is a human bacterial pathogen that multiplies only within an intracellular membrane-bound vacuole, the inclusion. *C. trachomatis* includes ocular and urogenital strains, usually causing infections restricted to epithelial cells of the conjunctiva and genital mucosa, respectively, and lymphogranuloma venereum (LGV) strains, which can infect macrophages and spread into lymph nodes. However, *C. trachomatis* genomes display >98% identity at the DNA level. In this Chapter, we studied whether *C. trachomatis* Inc proteins, which have a bi-lobed hydrophobic domain that might mediate their insertion in the inclusion membrane, could be a factor determining these different types of infection and tropism. Analyses of polymorphisms and phylogeny of 48 Inc proteins from 51 strains encompassing the three disease groups showed significant amino acid differences that were mainly due to variations between Inc proteins from LGV and ocular or urogenital isolates. Studies of the evolutionary dynamics of *inc* genes suggested that 10 of them are likely under positive selection and indicated that most non silent mutations are LGV specific. Additionally, real-time quantitative PCR analyses in prototype and clinical strains covering the three disease groups identified three *inc* genes with LGV-specific expression. We determined the transcriptional start sites of these genes and found LGV-specific nucleotides within their promoters. Thus, subtle variations in the amino acids of a subset of Inc proteins and in the expression of *inc* genes may contribute to the unique tropism and invasiveness of *C. trachomatis* LGV strains.

3.2 Introduction

Chlamydia trachomatis is a human obligate intracellular pathogen that includes the trachoma and LGV (lymphogranuloma venereum) biovars (Schachter, 1999). The trachoma biovar comprises ocular and urogenital strains, which cause localized infections of the epithelial surface of the conjunctiva and genital mucosa, respectively; strains of the LGV biovar cause invasive urogenital disease, due to their ability to infect macrophages and spread into lymph nodes. *C. trachomatis* strains can be further classified into ocular serovars A to C, urogenital serovars D to K, and LGV serovars L1 to L3.

The genomic sequences of different ocular, urogenital, and LGV strains exhibit >98% identity and a high degree of synteny (Stephens *et al.*, 1998; Carlson *et al.*, 2005; Thomson *et al.*, 2008; Seth-Smith *et al.*, 2009; Jeffrey *et al.*, 2010; Unemo *et al.*, 2010; Somboonna *et al.*, 2011; Harris *et al.*, 2012). Therefore, the determinants of the different types of infection (invasive or noninvasive) and tissue tropism (eyes, genitals, and lymph nodes) must rely on the few genes present in some strains but not in others and on nucleotide differences which could lead either to proteins with disease group-specific amino acids or to differential gene expression. Some of these determinants were suggested in previous studies: the tryptophan (*trpRBA*) operon (Allan C. Shaw *et al.*, 2000; Fehlner-Gardiner *et al.*, 2002; Caldwell *et al.*, 2003) and genes encoding cytotoxin (Carlson *et al.*, 2004), phospholipase (Nelson *et al.*, 2006), polymorphic membrane proteins (Pmps) (Gomes *et al.*, 2006), and Tarp (Lutter *et al.*, 2010).

Throughout development, *C. trachomatis* reside and multiply within a membranaceous compartment, known as the inclusion, and manipulate host cells by using a type III secretion system (T3SS) to translocate effector proteins into host cells (Valdivia, 2008; Betts *et al.*, 2009). These

effectors include the inclusion membrane (Inc) proteins, characterized by a bi-lobed hydrophobic motif thought to mediate their insertion into the inclusion membrane (Fields *et al.*, 2003; Subtil *et al.*, 2005). In Chapter 2, we found T3S signals in the majority of Inc proteins, supporting the notion that they are indeed T3S substrates. Inc proteins from the same chlamydial species are normally unrelated to each other, and only a subset of ~25 Inc proteins is conserved between species (Dehoux *et al.*, 2011; Lutter *et al.*, 2012). The *C. trachomatis* Inc proteins have been shown or suggested to subvert host cell vesicular and non-vesicular transport (Hackstadt *et al.*, 1999; Rzomp *et al.*, 2006; Delevoye *et al.*, 2008; Derré *et al.*, 2011). More recently, a few Incs were shown to be involved in the exit of the host cell by the bacteria (Lutter *et al.*, 2013), cytoskeleton reorganization (Dumoux *et al.*, 2015; Kokes *et al.*, 2015; Mital *et al.*, 2015), and disruption of endosomal trafficking (Mirrashidi *et al.*, 2015). Recently, it was also hypothesized that Inc proteins could interact with each other and act as scaffolds to form specific multi-protein complexes in the inclusion membrane (Mital *et al.*, 2013; Gauliard *et al.*, 2015). However, virtually nothing is known about the biological role of most Inc proteins.

In this Chapter, we used phylogenetic, molecular evolution, and gene expression analyses to determine whether Inc proteins could affect the type of infection and tissue tropism associated with *C. trachomatis*. Our studies suggest that a subset of Inc proteins might play a role in the unique capacity of *C. trachomatis* LGV strains to infect macrophages and disseminate into lymph nodes.

3.3 Materials and Methods

3.3.1 Bacterial strains and growth conditions

C. trachomatis prototype strains B/Har36, C/TW3, E/Bour, L2/434, and L3/404 [from the American Type Culture Collection (ATCC)] and clinical strains F/CS465-95 and L2b/CS19-08 (from the collection of the Portuguese National Institute of Health) were used as detailed below and propagated in HeLa 229 cells (from the ATCC) using standard techniques (Scidmore, 2005).

3.3.2 DNA sequences of C. trachomatis inc*,* pmp *and housekeeping genes*

Regardless of the origin of the *C. trachomatis* gene, throughout this work we used the nomenclature of the annotated D/UW3 strain (see Table A.3 in Annexes). The nucleotide sequences of the *inc*, *pmp*, and housekeeping genes analyzed were from the available genomes of 51 *C. trachomatis* strains (Stephens *et al.*, 1998; Carlson *et al.*, 2005; Thomson *et al.*, 2008; Seth-Smith *et al.*, 2009; Jeffrey *et al.*, 2010; Unemo *et al.*, 2010; Somboonna *et al.*, 2011; Harris *et al.*, 2012). The strains and corresponding genome accession numbers are listed in Table A.3 (Annexes). The DNA sequences were retrieved from pairwise alignments obtained by BLAST. All sequences were manually inspected and corrected for accuracy and completeness. We excluded from further analysis a few DNA sequences containing ambiguous nucleotides. Whenever there were distinct annotations in GenBank for the start codon of the same *inc* gene in *C. trachomatis* archetype ocular (A/Har13), urogenital (D/UW3), and LGV (L2/434) strains, we used the following: for *ct036* and *ct119* (*incA*), in all cases the start codons of each gene as annotated for D/UW3 and A/Har13; for *ct115* (*incD*) and *ct192*, the start codons of each gene as annotated for A/Har13, D/UW3,

or L2/434, in ocular, urogenital or LGV strains, respectively; for *ct226*, in all cases the start codon as annotated for A/Har13 (see Table 3.1 and Table A.3 in Annexes).

3.3.3 Sequence alignments and analyses of polymorphisms, phylogeny, and molecular evolution

Alignments of the amino acid sequences of the Inc, Pmp, and housekeeping proteins, deduced from the retrieved nucleotide sequences, were generated using the ClustalW algorithm in MEGA5 (Tamura *et al.*, 2011). All alignments were manually inspected and corrected for artifacts. We excluded all strain-specific pseudogenes (Table 3.1) from further phylogenomic and evolutionary analyses. For the analyses of polymorphism, phylogeny, and molecular evolution, various tools present in MEGA5 were used, essentially as previously described (Gomes *et al.*, 2006). Briefly, for analyses of polymorphism, we computed pairwise, overall, within-group (ocular, urogenital, or LGV), and between-group (ocular versus urogenital, ocular versus LGV, or urogenital versus LGV) amino acid *p* distances. For analyses of phylogeny, trees were generated using the neighbor-joining method (Saitou and Nei, 1987). The generated phylograms of Inc proteins were inspected for separate branches (segregation) of all ocular, urogenital, and LGV strains. This analysis was supported by comparison of pairwise amino acid *p* distances between and within disease groups. The phylogeny of an Inc protein was considered to segregate a disease group if the maximum pairwise amino acid *p* distance within groups was less than any of the pairwise amino acid *p* distances between groups. For molecular evolution analyses, we used the Kumar method (Nei and Kumar, 2000) to compute overall means of nonsynonymous (d_N) and synonymous (d_S) substitutions (per nonsynonymous or synonymous site, respectively) and to find genes that may be under positive selection. By

using the codon-based Z test of selection in MEGA5, genes were considered under positive selection if they showed a statistically significant value ($P < 0.05$) to reject the null hypothesis of strict neutrality ($d_N = d_S$) in favor of both positive selection ($d_N > d_S$) and lack of neutrality ($d_N \neq d_S$). All these analyses were performed by bootstrapping with 1,000 replicates and with the pairwise deletion option selected.

3.3.4 Real-time quantitative PCR

The expression of *inc* genes during the developmental cycle of *C. trachomatis* B/Har36, C/TW3, E/Bour, F/CS465-95, L2/434, L2b/CS19-08, and L3/404 was estimated by determining *inc* mRNA levels at different times postinfection by real-time quantitative PCR (RT-qPCR). These experiments were essentially done as previously described (Nunes *et al.*, 2007; Borges *et al.*, 2010). Briefly, for each strain, six tissue culture flasks with a surface area of 25 cm^2 containing monolayers of HeLa 229 cells were inoculated at a multiplicity of infection of 1; cells were harvested at 2, 6, 12, 20, 30, and 42 h postinfection by scraping in ice-cold phosphate-buffered saline. The cell suspension was sonicated to disrupt mammalian cells and promote bacterial release, followed by low-speed centrifugation at 4°C. The supernatant was then frozen in liquid nitrogen and stored at -80°C. These samples were used for total RNA purification and generation of cDNA, as previously described (Borges *et al.*, 2010). Primers were designed for each *inc* gene using Primer Express (Applied Biosystems), based on identical *C. trachomatis* sequences between strains (Table A.1 in Annexes). The RT-qPCR assays were done using the ABI 7000 SDS, SYBR green chemistry, and optical plates (Applied Biosystems), as previously described (Nunes *et al.*, 2007; Borges *et al.*, 2010). At each time point, raw RT-qPCR data for each *inc* gene were normalized against the data obtained for the 16S rRNA transcript, as it was

previously demonstrated that this is a good endogenous control (Borges *et al.*, 2010). The final results were based on at least two independent experiments.

3.3.5 Transcription linkage and identification of transcriptional start sites

We searched for disease group-specific nucleotides within the promoter region of three *inc* genes (*ct058*, *ct192*, and *ct214*) that showed differential gene expression. For this, we determined their transcriptional start sites (TSSs). For *ct058* and *ct192*, it was ambiguous whether their promoters would lie immediately upstream from their predicted start codons. Therefore, we used reverse transcription coupled with PCR (RT-PCR) to determine if *ct058* and *ct059*, or *ct192* and *ct193*, are part of the same transcriptional unit. For this, RNA was isolated from HeLa 229 cells infected for 30 h with *C. trachomatis* L2/434 using an NZY total RNA kit (NZYTech). cDNA was then generated by using random hexamers and iSCRIPT (Bio-Rad). Primers were designed to generate PCR products containing ~300 bp upstream and downstream from the predicted start codons (Table A.1 in Annexes). PCR products were obtained using DreamTaq DNA polymerase (Life Technologies). As controls for the PCRs, we also used as the template the product of a typical reverse transcription reaction but without iSCRIPT or total DNA isolated using an NZY tissue gDNA kit (NZYTech), from cells that were either infected with strain L2/434 for 42 h or left uninfected. The identification of the TSSs of *ct059* (upstream from *ct058* and in the same transcriptional unit), *ct192*, and *ct214* in L2/434 was done by 5' rapid amplification of cDNA ends (RACE), using a 5'/3' RACE kit (second generation; Roche). We used RNA isolated as described above from HeLa 229 cells infected for 30 h with *C. trachomatis* L2/434 and the primers listed in Table A.1 (Annexes). Final PCR amplification of double-

stranded cDNA was done with Phusion DNA polymerase (Life Technologies). PCR products were purified after agarose gel electrophoresis, using a High Pure PCR purification kit (Roche), and then subjected to DNA sequencing (Stab Vida). To analyze the determined TSSs in the context of the promoter regions of *ct059-ct058*, *ct192*, and *ct214* in all strains used in the RT-qPCR assays, the corresponding nucleotide sequences were either retrieved from GenBank (E/Bour, L2/434, and L3/404; Table A.3 in Annexes) or determined by DNA sequencing (C/TW3, B/Har36, F/CS465-95, and L2b/CS19-08), as previously described (Gomes *et al.*, 2006) and using the primers listed in Table A.1 (Annexes). All these manipulations were done according to instructions from the indicated manufacturers.

3.3.6 Nucleotide sequence accession numbers

The sequences of the promoter regions of *ct192* and *ct214* and of *ct059-ct058* in C/TW3, B/Har36, F/CS465-95, and L2b/CS19-08 determined in this study were submitted to GenBank and are available under accession numbers JX451863 to JX451874.

3.4 Results

3.4.1 Differences in the amino acid sequences of Inc proteins among C. trachomatis *strains correlate with the type of infection and with tissue tropism*

As in Chapter 2 we have seen that lack of a T3S signal in a protein containing a bi-lobed hydrophobic motif does not exclude its possible localization at the inclusion membrane, in this Chapter we focused on the complete set of 48 known and putative Inc proteins that was analyzed in Chapter 2 (Table 2.1 in Chapter 2). The nucleotide sequences of the genes encoding the selected proteins were retrieved from 51 fully sequenced *C. trachomatis* genomes (7 ocular, 23 urogenital, and 21 LGV strains) (Table A.3 in Annexes). In an initial analysis, 7 *inc* genes proved to be pseudogenes in different *C. trachomatis* strains (*ct058*, *ct101*, *ct135*, *ct192*, *ct227*, *ct228*, and *ct300*) (Table 3.1). *ct358* was described as a pseudogene in L2/434 and L2/UCH-1 (Thomson *et al.*, 2008), and its nucleotide sequence is 100% identical in all LGV strains. However, analysis of its nucleotide sequence suggests that *ct358* may encode a functional protein that is 7 amino acids shorter at its C-terminus than CT358 (178 amino acid residues) in ocular and urogenital strains (Table 3.1). In the seven *inc* genes that we identified as pseudogenes, the full-length gene is disrupted by a mutation that leads to a significantly truncated protein. The only disease group-specific correlation was observed with *ct300*, which is a pseudogene in all LGV strains analyzed. Therefore, the encoded protein is expendable for LGV infections. In addition, *ct058* and *ct101* revealed to be pseudogenes in almost all analyzed ocular and urogenital strains, respectively (Table 3.1). *ct135*, *ct192*, *ct227*, and *ct228*, are only

Table 3.1 – Deletion and insertion events, and pseudogenes in *inc* genes of *C. trachomatis*[a].

inc	**Deletions**			**Insertions**			**Pseudogenes**
gene	**position (nt)[a]**	**Nr. of codons**	**Strains**	**position (nt)[a]**	**Nr. of codons**	**Strains**	**strains**
ct058	333	1	A/Har13	123	1	A/Har13	Ocular except A/Har13
	834	4	LGV except L1/440				
ct101							Urogenital except D/UW3, E/11023, and E/Bour
ct115[b]				before start codon	10	Ocular	
				before start codon	14	LGV	
				411	4	LGV	
				423	1	LGV	
ct135				197	1	E/SW3	A/2497, A/Har13, B/Jali20, F/SW4, and J/6276
ct147	279	1	A/363, A/2497, and A/5291				
ct192[c]	before start codon	26	LGV				B/Jali20[c]
	before start codon	67	Ocular				
	210	8	D/SotonD1 and all E and F strains				
ct214				1110	1	Ocular	
ct222	222	1	LGV				
ct223	399	1	LGV				
	786	1	LGV				
ct226[d]	21	2	Ocular				
ct227							A/Har13
ct228							B/Jali20, B/TZ1A828/OT, D/SotonD6, and L1/115
ct288				1419	1	LGV	
ct300							LGV
ct358[e]	477	1	B/Jali20				
	492	3	LGV				
	519	5	LGV				

inc **gene**	**Deletions**			**Insertions**			**Pseudogenes**
	position (nt)[a]	**Nr. of codons**	**Strains**	**position (nt)[a]**	**Nr. of codons**	**Strains**	**strains**
ct383	717	4	A/363 and A/7249				
ct442	162	1	LGV	66	2	All except D/SotonD5, D/SotonD6, and G, Ia, J, and K strains	
	213	1	LGV, E and F strains, and D/SotonD1				

[a]The position is based in the relative *inc* gene sequence of *C. trachomatis* D/UW3. Whenever different start codons for the same *inc* gene were annotated in different strains, we followed the annotations from archetype ocular (A/Har13), urogenital (D/UW3), and LGV strains (L2/434) (except *ct226* and *ct358*, as indicated).

[b]Considering that *ct115* (*incD*) from urogenital strains have a different start codon from ocular and LGV strains, as annotated for A/Har13, D/UW3, and L2/434; however, if the start codon of *ct115* is identical for all strains and corresponds to the one annotated for *C. trachomatis* D/UW3 [as suggested by an upstream strong putative ribosome binding site (GUGAGG-N_3-AUG; predicted start codon underlined) that is present in all strains in the same position], then *ct115* is a pseudogene in ocular strains.

[c]Considering that *ct192* from ocular, urogenital and LGV strains have a different start codon, as annotated for A/Har13, D/UW3, and L2/434; however, similarly to *ct115*/*incD*, if the start codon of *ct192* is identical for all strains and corresponds to the one annotated for *C. trachomatis* L2/434/Bu [as suggested by an upstream strong putative ribosome binding site (AGGAGG-N_5-GUG; predicted start codon underlined) that is present in all strains in the same position], then *ct192* is a pseudogene in ocular strains. This is the predicted start codon of *ct192* annotated in A/2497, B/Jali20, and B/TZ1A428/OT (Seth-Smith *et al.*, 2009), and *ct192* is indicated as a pseudogene in these ocular strains (Seth-Smith *et al.*, 2009).

[d]Considering that the start codon of *ct226* is in all cases the one annotated for strain A/Har13. Otherwise, assuming the start codons annotated for strains D/UW3 and L2/434 for urogenital and LGV strains, then ocular strains have a 5 codon- and LGV strains an 11 codon-deletion relative to the start codon of urogenital strains, and *ct226* is a pseudogene in strain D/SotonD6.

[e]*ct358* is annotated as a pseudogene in L2/434 and its nucleotide sequence is 100% identical among LGV strains. However it is likely that in LGV strains this gene encodes a protein that is 7 amino acid residues shorter at its C-terminus than CT358 in ocular and urogenital strains [as annotated for LGV strain L2c (Somboonna *et al.*, 2011)].

pseudogenes in a few strains, with no obvious correlation with ocular, urogenital, or LGV disease groups (Table 3.1). Furthermore, 12 *inc* genes showed small deletion and insertion events (Table 3.1).

To understand if the amino acid sequences of Inc proteins vary among strains, we determined the overall mean genetic distance (amino acid *p* distance) for each Inc protein among all 51 *C. trachomatis* strains (discarding strain-specific pseudogenes) (Figure 3.1A). As a reference, we also analyzed the 9 Pmps and 9 housekeeping proteins of *C. trachomatis* previously shown to be polymorphic (Nunes *et al.*, 2008) (Figure 3.1A). The Pmps should localize at the bacterial outer membrane and the housekeeping proteins within the bacterial cell. The average *p* distance was 0.017 (SEM, 0.002) for Inc proteins, 0.020 (SEM, 0.007) for Pmps, and 0.013 (SEM, 0.003) for housekeeping proteins. Based on this, and considering the average *p* distance for Inc proteins as a cutoff value, we defined 19 (40% of the total) Inc proteins as polymorphic (*p* distance ≥ 0.017) (Figure 3.1A). As comparison, 4 Pmps (44%) and 2 housekeeping proteins (22%) displayed a *p* distance of ≥ 0.017. This showed that the overall degree of polymorphism in Inc proteins among *C. trachomatis* strains is similar to that of Pmps and is higher than that of known polymorphic housekeeping proteins.

To understand if the amino acid differences between Inc proteins were related to the type of infection and with tissue tropism, we determined the average *p* distances of Inc proteins within and between the three groups of strains (ocular, urogenital, and LGV) (Figure 3.1B). This showed that the differences were largely due to variations between Inc proteins from LGV strains and ocular (average *p* distance = 0.031; SEM, 0.004) or urogenital strains (average *p* distance = 0.029; SEM, 0.004) (Figure 3.1B). These average *p* distances were significantly higher (in all cases, $P < 0.0001$; two-tailed *t* test) than those between Inc proteins

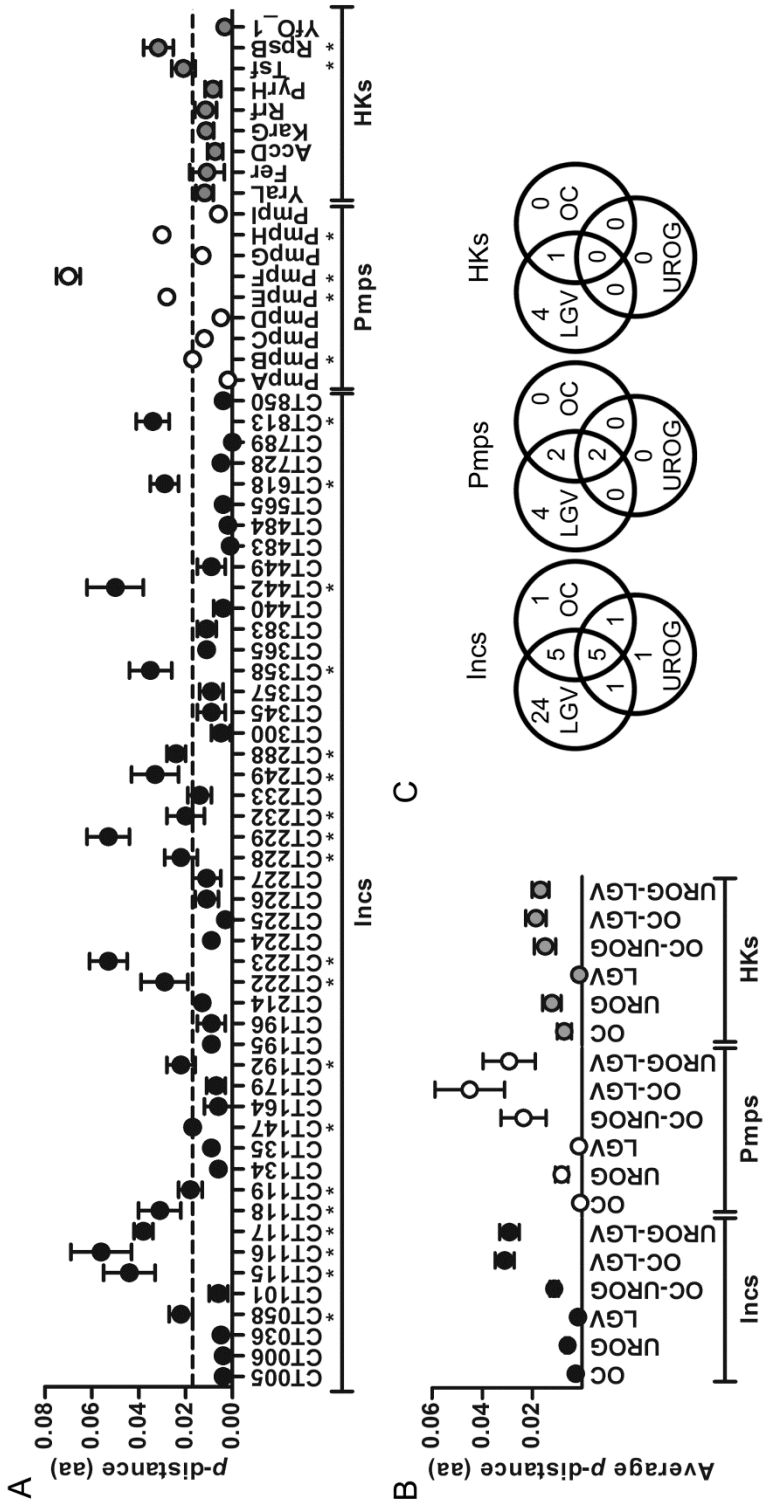

A
p-distance (aa)
0.08
0.06
0.04
0.02
0.00
CT005
CT006
CT036
*CT058
CT101
*CT115
*CT116
*CT117
*CT118
*CT119
CT134
CT135
*CT147
CT164
CT179
*CT192
CT195
CT196
CT214
*CT222
*CT223
CT224
CT225
CT226
CT227
*CT228
*CT229
*CT232
CT233
*CT249
*CT288
CT300
CT345
CT357
*CT358
CT365
CT383
CT440
*CT442
CT449
CT483
CT484
CT565
*CT618
CT728
CT789
*CT813
CT850
PmpA
*PmpB
PmpC
PmpD
*PmpE
*PmpF
PmpG
*PmpH
PmpI
YraL
Fer
AccD
KarG
Rrf
PyrH
Tsf
*RpsB
YfiO_1
Incs
Pmps
HKs
B
Average p-distance (aa)
0.06
0.04
0.02
OC
UROG
LGV
OC-UROG
OC-LGV
UROG-LGV
OC
UROG
LGV
OC-UROG
OC-LGV
UROG-LGV
OC
UROG
LGV
OC-UROG
OC-LGV
UROG-LGV
Incs
Pmps
HKs
C
Incs
24 LGV
5 OC
1 UROG
1
5
1
Pmps
4 LGV
0 OC
0 UROG
2
0
2
0
HKs
4 LGV
0 OC
0 UROG
1
0
0
0
0

Figure 3.1 – Polymorphisms in *C. trachomatis* Inc proteins. Polymorphic membrane proteins (Pmps) and housekeeping proteins (HKs) were analyzed as references. (A) Overall mean genetic distance (polymorphisms) based on the *p* distance between all possible pairs of amino acid (aa) sequences of Inc proteins, Pmps, and HKs among *C. trachomatis* strains. Proteins marked with an asterisk have a *p* distance that is equal to or higher than the average value for Inc proteins [0.017 (dashed line)]. (B) Average mean genetic distance, based on the *p* distance between all possible pairs of amino acid sequences within [ocular (OC), urogenital (UROG), or LGV] or between (OC-UROG, OC-LGV, or UROG-LGV) *C. trachomatis* disease groups. (C) Venn diagrams showing the phylogenetic segregation of *C. trachomatis* disease groups based on neighbor-joining trees of Inc proteins, Pmps, or HKs and on pairwise *p* distances between and within disease groups for all possible pairs of Inc protein, Pmp, or HK sequences from the *C. trachomatis* strains analyzed. All these analyses were performed by bootstrapping with 1,000 replicates. Error bars represent SEM.

from ocular and urogenital strains (average *p* distance = 0.011; SEM, 0.002) or within Inc proteins from the same disease groups (average *p* distance < 0.006) (Figure 3.1B). Similar observations were made for Pmps except that, in contrast to Inc proteins, the average *p* distance between Pmps from ocular and urogenital strains was not significantly different than that between Pmps from LGV strains and ocular or urogenital strains (in both cases, $P > 0.05$; two-tailed *t* test) (Figure 3.1B). Also in contrast to the Inc proteins, the variation in the amino acid sequences of housekeeping proteins between each of the three groups was clearly not different (in all cases, $P > 0.05$; two-tailed *t* test) (Figure 3.1B). Furthermore, the amino acid sequences of housekeeping proteins varied nearly as much between or within groups (Figure 3.1B). The main exception was the LGV group, within which the housekeeping proteins, like Inc proteins and Pmps, proved to be extremely highly conserved (average *p* distance < 0.002 for the three cases) (Figure 3.1B). Thus, the separation (*p* distance) between LGV strains and ocular or urogenital strains is much more marked for Inc proteins than for Pmps or housekeeping proteins.

We then made and analyzed phylogenetic reconstructions based in the amino acid sequence of Inc proteins. The phylograms of five Inc proteins (10%) showed tropism, i.e., segregation of the three disease groups, and those of 38 Inc proteins (84%) evidenced segregation of at least one disease group (Table 3.2 and Figure 3.1C). The phylograms of a total of 35 Inc proteins (73%) showed segregation of LGV strains, while only 12 (25%) and 8 (17%) displayed clustering of ocular and urogenital strains, respectively (Table 3.2 and Figure 3.1C). This scenario was mirrored by the phylograms of Pmps but was in contrast to the phylograms of housekeeping proteins, in which segregation by disease group was less often seen (Table 3.2 and Figure 3.1C).

Table 3.2 – Distribution of the segregation into disease groups displayed by amino acid-based phylograms of Incs, Pmps, and housekeeping proteins (HKs).

Segregation	Incs	Pmps	HKs
Tropism	CT115, CT135, CT214, CT223, CT383	PmpF, PmpH	
Ocular & Urogenital	CT232		
Ocular & LGV	CT116, CT195, CT224, CT442, CT850	PmpE, PmpI	YraL
LGV & Urogenital	CT226		
Ocular	CT225		
Urogenital	CT233		
LGV	CT006, CT036, CT101, CT117, CT118, CT147, CT179, CT192, CT196, CT222, CT227, CT228, CT229, CT249, CT288, CT345, CT357, CT358, CT365, CT449, CT484, CT565, CT618, CT813	PmpB, PmpC, PmpD, PmpG	Fer, AccD, PyrH, Rs2
None	CT005, CT058, CT119, CT134, CT164, CT300, CT440, CT483, CT728, CT789	PmpA	KarG, Frr, Tsf, YfO_1

In summary, there are significant differences in the amino acid sequences of Inc proteins among *C. trachomatis* strains, and they correlate with the type of infection and with tissue tropism. In particular, differences are mostly between Inc proteins from LGV strains and Inc proteins from ocular or urogenital strains.

3.4.2 inc *genes of* C. trachomatis *have distinct evolutionary dynamics, and several* inc *genes are likely under positive selection*

To understand the underlying evolutionary pressures that drive amino acid changes in Inc proteins, we analyzed the molecular evolution of *inc* genes. We first determined overall d_N/d_S values for *inc* genes by comparing them to the 9 *pmp* genes and the 9 selected housekeeping genes of *C. trachomatis*. We found that 24 *inc* genes (50%) had d_N/d_S values of >1 and that in four *inc* genes, all substitutions were nonsynonymous (Figure 3.2A and Table A.4 in Annexes). In contrast, only two *pmp* genes (22%) had d_N/d_S values of >1, and all housekeeping genes had d_N/d_S values of <1 (Figure 3.2A and Table A.4 in Annexes).

Analyses of the overall d_N/d_S values using the codon-based Z test of selection (see Materials and Methods) yielded statistically significant values for 10 *inc* genes (*ct116*/*incE*, *ct118*/*incG*, *ct119*/*incA*, *ct222*, *ct223*, *ct228*, *ct229*, *ct249*, *ct288*, and *ct813*) but only for one *pmp* gene and for no housekeeping gene (Figure 3.2A and Table A.4 in Annexes). This indicated that these 10 *inc* genes are likely under positive selection. We then aimed to understand which group of strains might cause the detection of this possible evolutionary trend. For this, we assessed the impact of artificially removing ocular, urogenital, or LGV strains from the analyses (Figure 3.2B and Tables A.4, A.5, A.6, and A.7 in Annexes).

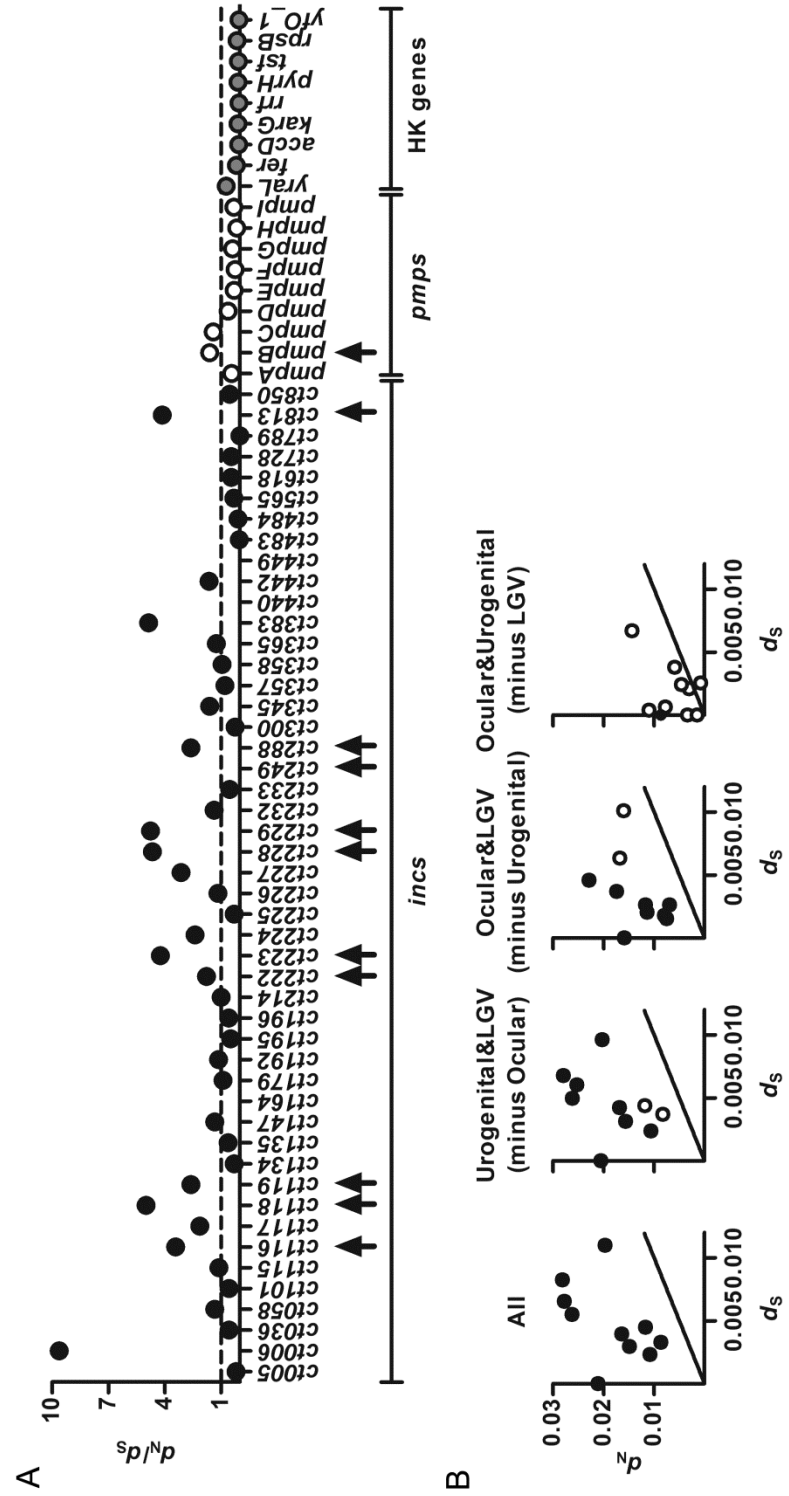

A
B
d_N/d_S
10
7
4
1
ct005
ct006
ct036
ct058
ct101
ct115
ct116
ct117
ct118
ct119
ct134
ct135
ct147
ct164
ct179
ct192
ct195
ct196
ct214
ct222
ct223
ct224
ct225
ct226
ct227
ct228
ct229
ct232
ct233
ct249
ct288
ct300
ct345
ct357
ct358
ct365
ct383
ct440
ct442
ct449
ct483
ct484
ct565
ct618
ct728
ct789
ct813
ct850
pmpA
pmpB
pmpC
pmpD
pmpE
pmpF
pmpG
pmpH
pmpI
yraL
fer
accD
karG
rrf
pyrH
tsf
rpsB
yfiO_1
incs
pmps
HK genes
All
Urogenital&LGV (minus Ocular)
Ocular&LGV (minus Urogenital)
Ocular&Urogenital (minus LGV)
p_N
0.03
0.02
0.01
0.0050.010
d_S

Figure 3.2 – Evolutionary dynamics of *inc* genes. Genes encoding polymorphic membrane proteins (*pmp* genes) or housekeeping (HK) proteins were used as reference. (A) Ratio of nonsynonymous (d_N) to synonymous (d_S) substitutions among *C. trachomatis* strains. The dashed line indicates neutrality ($d_N/d_S = 1$). The arrows specify genes likely under positive selection, according to the codon-based Z test of selection (see Materials and Methods). (B) Distribution of d_N versus d_S values for the 10 *inc* genes likely under positive selection, comparing the impact of artificially discarding ocular, urogenital, or LGV strains relative to the analysis with all *C. trachomatis* strains. The straight line in each graph indicates neutrality ($d_N/d_S = 1$). In all cases, *inc* genes likely under positive selection (codon-based Z test of selection) are depicted as black circles, while *inc* genes for which statistical support of likely positive selection can no longer be detected after discarding a particular group of strains are depicted as white circles. All these analyses were performed by bootstrapping with 1,000 replicates. For sake of clarity, SEMs for values in both panels are presented only in Tables A.4, A.5, A.6, and A.7 in Annexes.

This showed that discarding LGV strains caused major alterations on the d_N and d_S values and confined the ability to detect Z-test-based likely positive selection to only one *inc* gene, whereas discarding ocular or urogenital strains had less pronounced effects (Figure 3.2B and Tables A.4, A.5, A.6, and A.7 in Annexes). In addition, within each of the 10 proteins encoded by *inc* genes likely under positive selection, we found 113 amino acid residues that are disease group specific (Table 3.3). Among these residues, 104 (92%) were in Inc proteins from LGV strains, 77 (74%) of which localized in regions of the proteins predicted to be on the cytoplasmic side of the inclusion membrane (Table 3.3). Overall, this suggested that the evolutionary dynamics of *inc* genes is distinct from that of *pmp* genes or housekeeping genes and that LGV-specific amino acid residues in a subset of Inc proteins might be involved in the unique ability of *C. trachomatis* LGV strains to infect macrophages and disseminate to lymph nodes.

Table 3.3 – Amino acids residues within Inc proteins encoded by genes likely under positive selection that are specific of *C. trachomatis* disease groups

Inc	Length (aa)[a]	TM segments[b]	LGV[c]	Urogenital[c]	Ocular[c]
CT116 (IncE)	132	[36-59], [64-87]	V37**A**, V38**A**, S39**C**, S64**G**, I68**L**, V72**I**, I84**T**, D92**N**	D19**G**	
CT118 (IncG)	167	[33-57], [63-88]	A33**V**, F77**C**, C79**Y**, N87**S**, F135**L**, G136**R**, H165**R**		
CT119 (IncA)	273	[35-59], [64-84]	I75**T**, Q207**E**, V211**A**, V212**A**		
CT222	129	[39-63], [69-93]	Y10**C**, I75**-**, C89**Y**, L125**I**, V126**S**, Y127**V**, S128**F**, N129**H**		
CT223	270	[38-61], [67-91]	G11**R**, A65**T**, K100**R**, I108**L**, K127**G**, N130**D**, P134**-**, C153**Y**, E160**D**, T166**K**, H200**Y**, E204**D**, N207**R**, L208**M**, R231**L**, V241**A**, P251**L**, D261**Y**, G263**-**	L225**F**	A40**V**, A75**V**, D142**G**, V161**M**
CT228	196	[38-59], [65-86]	I22**T**, A68**V**, A72**P**, C90**Y**, V119**A**, A143**V**, I144**M**, V146**F**		
CT229	215	[42-65], [71-90]	A39**S**, V99**I**, G112**E**, K114**E**, F121**S**, Q125**R**, V126**A**, H137**Y**, Q144**K**, Y152**H**, E154**A**, E164**K**, G173**R**, S181**N**, T201**A**	Q102**H**, I156**V**	C157**S**
CT249	116	[51-72], [78-97]	H8**Y**, D24**N**, V80**T**, A89**I**		
CT288	563	[36-58], [65-88], [242-263], [269-291]	A67**T**, I69**V**, E102**A**, S121**P**, A194**T**, S198**T**, N203**K**, T209**I**, A258**T**, I260**V**, I261**V**, L318**W**, S346**G**, L388**V**, -474**D**, A516**P**, V531**L**, A532**T**		
CT813	264	[41-61], [68-94]	N7**T**, V76**I**, E107**K**, K125**E**, R128**Q**, T145**A**, E161**K**, A163**G**, E168**K**, E171**K**, E172**Q**, I181**V**, I236**T**		

[a]Based on the protein sequence annotation of *C. trachomatis* strain D/UW3.

[b]Positions of transmembrane (TM) domains in the corresponding Inc proteins, obtained from reference (Dehoux *et al.*, 2011), except for CT223, for which we found only the two indicated TM domains.

[c]Residues specific to Inc proteins from each disease group are in bold, relative to the amino acid in the same position in the other disease groups, and those in regions predicted to be on the cytoplasmic side of the inclusion membrane are underlined (we considered that the loop region within two TM segments faces the lumen of the inclusion); a dash indicates a deletion or an insertion.

3.4.3 Disease group-specific expression of C. trachomatis inc *genes*

To analyze if there were differences in the expression of *inc* genes between *C. trachomatis* strains that correlate with the type of infection or with tissue tropism, we used RT-qPCR to determine the mRNA levels of the 48 selected *inc* genes (Table 2.1 in Chapter 2) throughout the developmental cycle of *C. trachomatis*. We aimed to find *inc* genes showing differences in the highest mRNA levels during the cycle (peak of expression) or in the variation of mRNA levels throughout development (profile of expression) between *C. trachomatis* strains.

We first infected HeLa 229 cells with ocular (C/TW3), urogenital (E/Bour), or LGV (L2/434) prototype strains. Total RNA was isolated at 2, 6, 12, 20, 30, and 42 h postinfection, which was used to generate cDNA for RT-qPCR assays (complete data are shown in Tables A.8, A.9, and A.10 in Annexes). Generally, the comparison of the peak of expression revealed differences from 30- to 60-fold (depending on the strain) between *inc* genes (Figure 3.3A). However, the averages of the peaks of expression of *inc* genes in C/TW3, E/Bour, and L2/434 were not significantly different between strains (Figure 3.3A). Regarding the profile of expression, essentially as previously described (E. I. Shaw *et al.*, 2000; Belland *et al.*, 2003; Nicholson *et al.*, 2003) we identified *inc* genes whose expression was highest at 2 or 6 h postinfection and then either decreased or remained constant throughout the cycle (early-cycle genes), *inc* genes whose expression was highest only at 12 or 20 h postinfection and then decreased or remained constant at later time points (mid-cycle genes), and *inc* genes whose expression was highest only at 30 or 42 h postinfection (late-cycle genes) (Figure 3.3B and C).

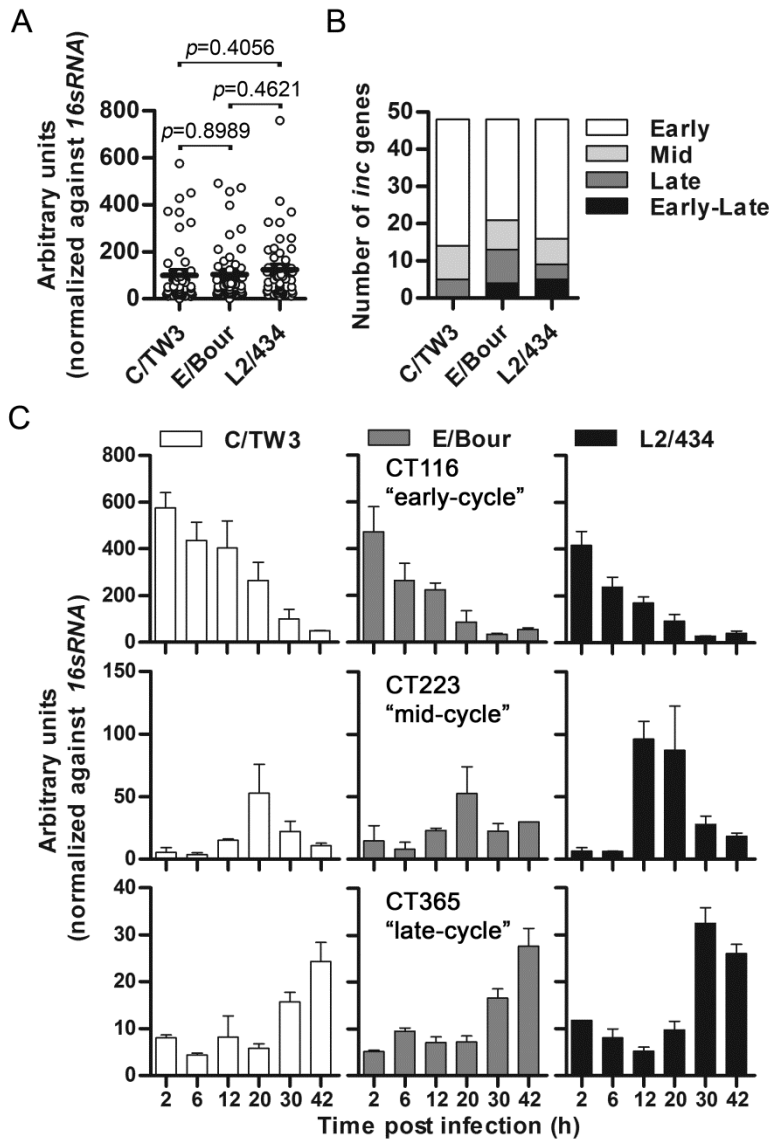


Figure 3.3 – mRNA levels of *inc* genes during the developmental cycle of different *C. trachomatis* strains. The mRNA levels of 48 *inc* genes were analyzed by RT-qPCR throughout the developmental cycle of the prototype strains C/TW3, E/Bour, and L2/434. (A) Peak of expression (highest mRNA levels during the developmental cycle) of each *inc* gene. The *P* values were calculated by two-tailed *t* tests. (B) Number of *inc* genes showing the indicated profiles of expression (variation of mRNA levels during the developmental cycle). (C) Examples of different profiles of expression. The expression values (mean ± SEM) resulted from raw RT-qPCR data (x10^5) of each gene normalized to that of the 16S rRNA gene and are from at least two independent experiments. Complete data are shown in Tables A.8, A.9, and A.10 (Annexes).

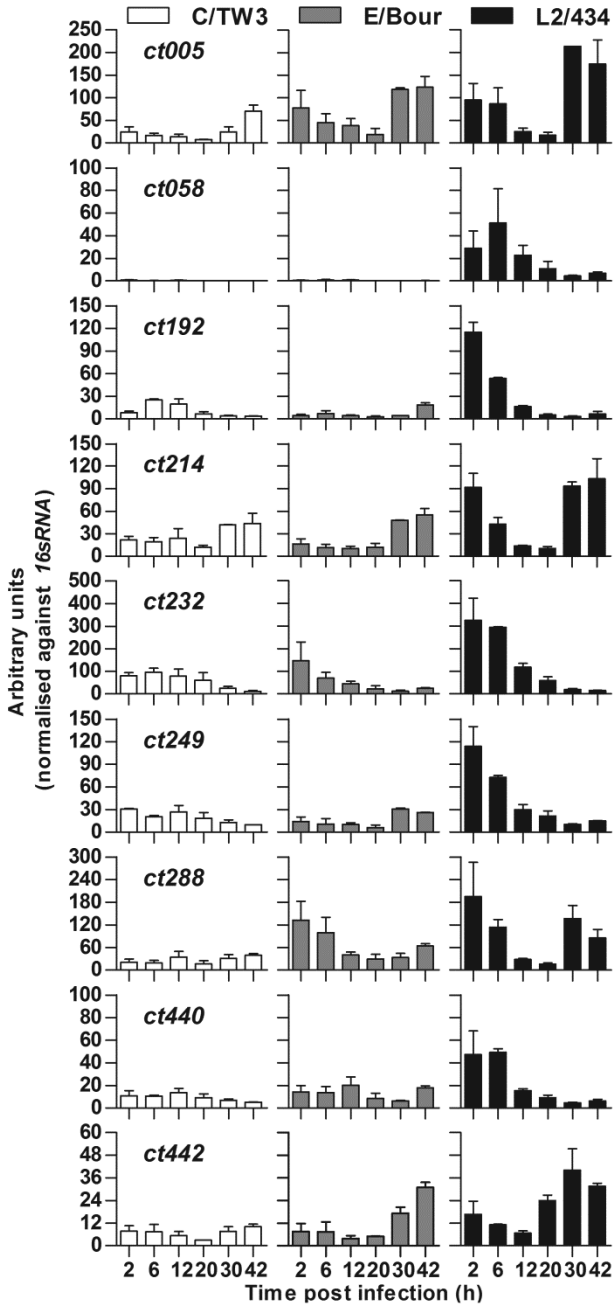


Figure 3.4 – Differences in the mRNA levels of *inc* genes throughout the developmental cycle of *C. trachomatis* C/TW3, E/Bour, and L2/434. The expression values (mean ± SEM) of the indicated inc genes result from the RT- qPCR raw data ($x10^5$) of each gene normalized against the 16SrRNA, from at least two independent experiments. Complete data are shown in Tables A.8, A.9, and A.10 (Annexes).

We also identified four *inc* genes in E/Bour and five *inc* genes in L2/434 that were simultaneously early- and late-cycle genes, showing identically high mRNA levels both at 2 and 6h postinfection and at 30 and 42h postinfection but lower expression at mid-cycle (Figure 3.3B). This was typically the case of *ct214* and *ct288* in L2/434 (Figure 3.4). Generally, 35 *inc* genes showed the same profile of expression in the three *C. trachomatis* strains, and the majority of *inc* genes showed an early-cycle profile of expression (34 in C/TW3, 27 in E/Bour, and 32 in L2/434) (Figure 3.3B). In spite of these common features, we identified 9 *inc* genes (*ct005*, *ct058*, *ct192*, *ct214*, *ct232*/*incB*, *ct249*, *ct288*, *ct440*, and *ct442*) whose peak of expression consistently showed >2-fold differences between strains and/or whose profile of expression displayed differences that could not be explained by distinct growth kinetics of the strains (Figure 3.4). With the exception of *ct214*, all these genes showed consistently higher peaks of expression in L2/434 than in C/TW3 and/or E/Bour; *ct442* also showed a higher peak of expression in E/Bour than in C/TW3 (Figure 3.4).

To analyze whether the differences found in the expression of *ct005*, *ct058*, *ct192*, *ct214*, *ct232*/*incB*, *ct249*, *ct288*, *ct440*, and *ct442* were strain specific or disease group specific, we determined the mRNA levels of these genes during the developmental cycle of additional *C. trachomatis* strains. For this, we infected HeLa 229 cells with ocular B/Har36 (prototype), urogenital F/CS465-95 (clinical isolate), LGV L2b/CS19-08 (clinical isolate), or LGV L3/404 (prototype) strains. The infected cells were processed for RT-qPCR assays and analyzed as described above (complete data are shown in Tables A.8, A.9, A.10, A.11, A.12, A.13, and A.14 in Annexes). We detected disease group-specific differences in gene expression only for *ct058*, *ct192*, and *ct214* (Figure 3.5): *ct058* showed an early-cycle gene profile of expression in which mRNA levels were evident for LGV strains but only vestigial for

ocular and urogenital strains; *ct192* showed only a clear early-cycle gene profile of expression in LGV strains, and its expression levels were generally higher in LGV strains than in ocular or urogenital strains; *ct214* displayed an early- and late-cycle gene profile of expression in LGV strains but a late-cycle gene profile of expression in ocular or urogenital strains.

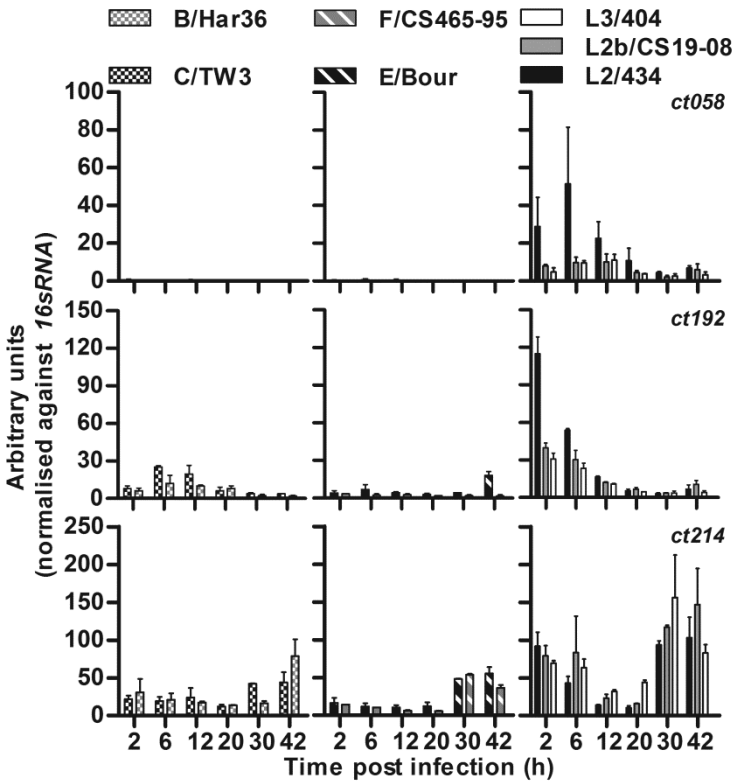


Figure 3.5 – Disease group-specific differences in the mRNA levels of *ct058*, *ct192*, and *ct214* throughout the *Chlamydia trachomatis* infectious cycle. The mRNA levels of *ct058*, *ct192*, and *ct214* were analyzed by RT-qPCR throughout the developmental cycle of the indicated prototype (B/Har36, C/TW3, E/Bour, L2/434, and L3/404) and clinical (F/CS465-95 and L2b/CS19-08) strains. The expression values (mean ± SEM) resulted from raw RT-qPCR data ($x10^5$) of each gene normalized to that of the 16S rRNA gene and are from at least two independent experiments. Complete data are shown in Tables A.8, A.9, A.10, A.11, A.12, A.13, and A.14 (Annexes).

Therefore, we have identified three *inc* genes (*ct058*, *ct192*, and *ct214*) with differences in expression between *C. trachomatis* strains that correlate with the type of infection and tissue tropism, in particular with LGV isolates.

3.4.4 Identification of LGV-specific nucleotides in the promoter regions of ct058*,* ct192*, and* ct214

We next attempted to obtain insights into the genetic basis for the disease group-specific expression of *ct058*, *ct192*, and *ct214* by analyzing the promoter regions of these genes in *C. trachomatis* L2/434. The gene organization of these *inc* genes suggested that the promoter region of *ct214* should lie between the start codons of *ct214* and *ct215* or within the first codons of *ct215* (Figure 3.6A). However, the localization of the promoter regions of *ct058* or *ct192* was unclear, as these genes could be co-transcribed with *ct059* or *ct193*, respectively (Figure 3.6A). Therefore, we used RT-PCR with a cDNA template generated from total RNA of HeLa 229 cells infected with *C. trachomatis* L2/434 to determine the possible transcriptional linkages between *ct058* and *ct059* and between *ct192* and *ct193*. This indicated that *ct058* is co-transcribed with *ct059*, which likely encodes a ferredoxin, and transcription of *ct192* is unlinked from *ct193* (Figure 3.6B). We previously detected clearly measurable mRNA levels of *ct059* in *C. trachomatis* ocular and urogenital strains (Borges *et al.*, 2010). We confirmed this using the same biological samples in which the levels of *ct058* mRNA were vestigial (data not shown). Therefore, we tentatively propose that expression of *ct058* in ocular and urogenital strains might be downregulated by specific 3'-to-5' posttranscriptional processing of the *ct059-ct058* transcript.

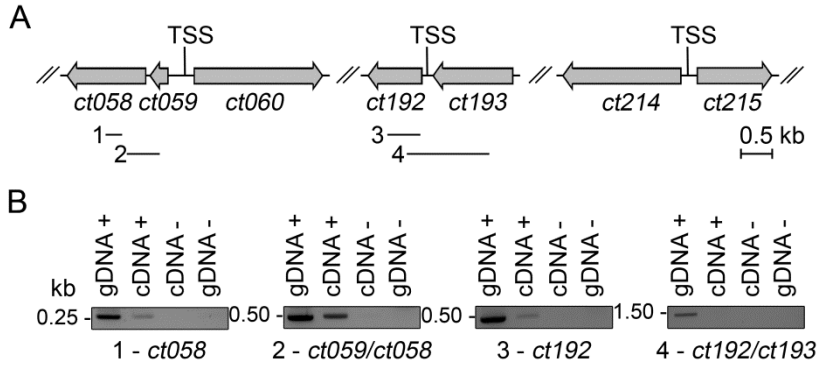


Figure 3.6 - Genetic organization of *ct058*, *ct192*, and *ct214*. (A) Scheme of the genetic organization of the three *inc* genes (the nomenclature of D/UW3 is used) depicting the fragments amplified in the transcriptional linkage analysis and the approximate locations of the transcriptional start sites (TSSs) determined by RACE in L2/434. (B) Transcriptional linkage analysis in L2/434: gDNA+, PCR from total DNA isolated from cells infected with L2/434; cDNA+, PCR from cDNA generated with reverse transcriptase (RT) from total RNA isolated from cells infected with L2/434; cDNA-, as for cDNA+ but without RT; gDNA-, PCR from DNA of uninfected cells.

To precisely define the promoter regions of *ct059-ct058*, *ct192*, and *ct214*, we determined their TSSs by RACE, using as the template total RNA of cells infected with *C. trachomatis* L2/434 and primers complementary to the *ct059*, *ct192*, or *ct214* mRNA (Figure 3.7). When we used a primer complementary to the *ct058* mRNA, we were unable to identify a *ct058*-exclusive TSS upstream from its start codon in L2/434 (data not shown). The TSS of *ct214* matched the one previously identified by deep sequencing in strain L2b/UCH-1 (Albrecht *et al.*, 2009), while the TSSs of *ct059-ct058* and *ct192* have not been identified before.

C. trachomatis encodes three σ factors: σ^{66}, the homolog of the *E. coli* main σ factor (σ^{70}); σ^{28}, a minor σ factor; and σ^{54}, an alternative σ factor (Mathews and Timms, 2006). By inspecting the nucleotide sequences immediately upstream from the determined TSSs of *ct059-ct058*, *ct192*,

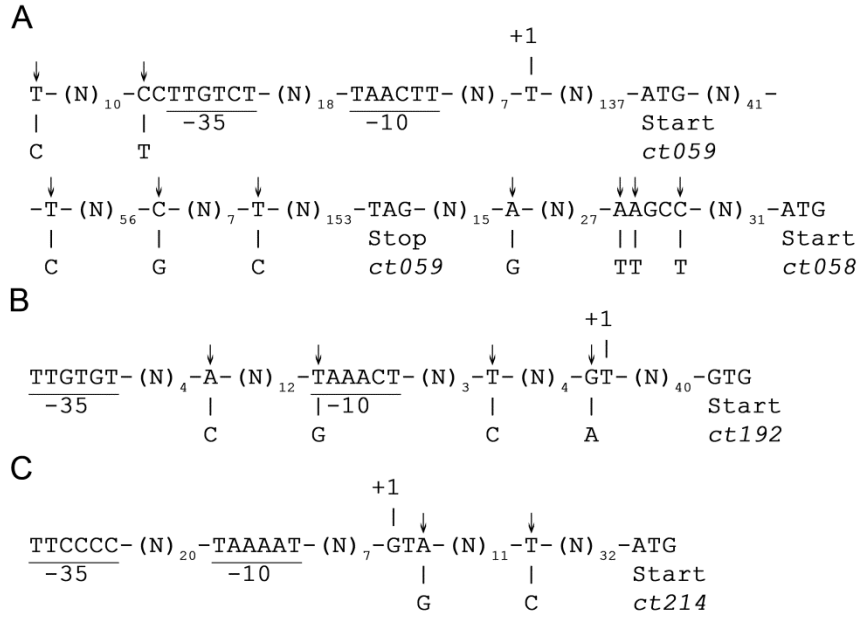


Figure 3.7 – Identification of LGV-specific nucleotides in the promoter regions of *ct192* and *ct214* and within *ct059-ct058*. Schematic view of the nucleotide sequences of the *ct059* promoter region, *ct059-ct058* intragenic region, and first codons of *ct058* (A), and of the promoter regions of *ct192* (B) and *ct214* (C) in L2/434 strain. The TSSs are labeled +1, and the predicted -10 and -35 σ^{66}-like hexamers are underlined. The sequences of the three LGV strains used in RT-qPCR assays are 100% identical within the depicted regions (Figure 3.5 and Figure A.1 in Annexes). The identified LGV-specific nucleotides are indicated with an arrow, and letters below the sequence represent nucleotides present in those positions in the ocular and urogenital strains used in RT-qPCR assays. The full nucleotide sequences of these regions in the strains used in RT-qPCR assays are shown in Figure A-1 (Annexes).

and *ct214* for σ^{66}-, σ^{54}-, and σ^{28}-like promoters (Mathews and Timms, 2006), we identified only σ^{66}-like promoters (Figure 3.7 and Figure A.1 in Annexes). Finally, we analyzed the promoter regions of *ct059-ct058*, *ct192*, and *ct214* and the *ct059-ct058* transcript for LGV-specific nucleotide differences in all strains used in the RT-qPCR assays. We found LGV-specific nucleotides upstream from the predicted -35 region of *ct059* (Figure 3.7A and Figure A.1A in Annexes). Furthermore, we found seven LGV- specific nucleotides scattered within the coding region

of *ct059* and between the stop codon of *ct059* and the start codon of *ct058* (Figure 3.7A and Figure A.1A in Annexes). In particular, three of these LGV-specific nucleotide differences were clustered 30 to 35 nucleotides upstream from the start codon of *ct058* (Figure 3.7A and Figure A.1A in Annexes). However, it is unclear how these LGV-specific differences could explain the vestigial mRNA levels of *ct058* in ocular and urogenital strains or have a discriminatory role in the proposed hypothetical degradation of the *ct058* transcript in those strains. The scenario was simpler for *ct192* and *ct214*, as we identified discrete LGV-specific nucleotides within the promoter regions of these genes that may explain their disease group-specific expression (Figure 3.7B and C and Figure A.1B and C in Annexes).

3.5 Discussion

In this Chapter, we found that amino acid differences between Inc proteins and distinct mRNA levels among *inc* genes throughout the developmental cycle of *C. trachomatis* strains correlate with the specific invasiveness and tropism of LGV isolates. Thus, we propose the novel hypothesis that a subset of Inc proteins may contribute to the specificity of infection by LGV strains. In fact, the vast majority of amino acid differences between Inc proteins are due to variations between proteins from LGV and trachoma biovars (Figure 3.1). This could simply reflect the evolutionary history of *C. trachomatis* (Harris *et al.*, 2012). However, most *inc* genes had d_N/d_S values of >1 among *C. trachomatis* strains, and according to the Z test of selection, 10 *inc* genes are likely under positive selection. In contrast, *pmp* genes or selected housekeeping genes of *C. trachomatis* mostly had d_N/d_S values of <1 (Figure 3.2). Moreover, polymorphisms in *C. trachomatis* genomes are essentially driven by fixation of silent mutations (Borges *et al.*, 2012). This suggests that the amino acid differences between Inc proteins should not be explained solely by genetic drift. In addition, almost all disease group-specific amino acids of Inc proteins encoded by genes likely under positive selection were found among proteins of LGV strains, and the majority of these amino acids localized in regions of the proteins predicted to face the host cell cytosol (Table 3.3). However, it must be clarified that other proteins in addition to Inc proteins are likely involved in the specificity of infection by LGV strains (Joseph *et al.*, 2011; Borges *et al.*, 2012). Moreover, the overall determinants of tissue tropism of *C. trachomatis* should be complex and multifactorial and should certainly also include, e.g., the products of the *trpRBA* operon (Allan C. Shaw *et al.*, 2000; Fehlner-Gardiner *et al.*, 2002; Caldwell *et al.*, 2003) or of the cytotoxin gene (Carlson *et al.*, 2004) and Tarp (Lutter *et al.*, 2010).

We have done a focused and in-depth study of the variability of Inc proteins and evolution and expression of *inc* genes among *C. trachomatis* serovars, encompassing most of the available genomic information. We reveal that the overall degree of variation in the amino acid sequences of Inc proteins among strains is similar to that of a characteristic family of *C. trachomatis* polymorphic proteins (Pmps). Our results strengthen previous studies that suggested that some *C. trachomatis* Inc proteins could contribute to tissue tropism (Brunelle *et al.*, 2004; Nunes *et al.*, 2008; Borges *et al.*, 2012; Lutter *et al.*, 2012) and confirm recent data suggesting that many *inc* genes could be under positive selection (Joseph *et al.*, 2011; Borges *et al.*, 2012). However, almost all of these previous studies analyzed a limited number of sequences, as the majority of the 51 genomic sequences used in this Chapter became available only very recently (Harris *et al.*, 2012).

Our findings also revealed that differential gene expression could be a mechanism contributing to the different invasiveness and tissue tropism of *C. trachomatis* strains. A previous work have identified differences in expression between *pmp* genes from reference and clinical strains (Nunes *et al.*, 2007), but disease group-specific differences in expression of *C. trachomatis* genes have not been noticed before. The *inc* genes *ct058*, *ct192*, and *ct214* evidenced LGV-specific gene expression (Figure 3.5), and we have further identified LGV-specific nucleotides in the promoter regions of *ct192* and *ct214* and within the *ct059-ct058* transcript (Figure 3.7). This was directly analyzed for the strains used in RT-qPCR assays, but the specificity is maintained within all 51 *C. trachomatis* genomes (Table A.3 in Annexes) that we used in our studies (data not shown). As only 3 of 48 *inc* genes showed LGV-specific gene expression, it is unlikely that this specificity is a common feature among *C. trachomatis* genes.

For *ct214*, we tentatively propose that the LGV-specific nucleotides could differentially affect its expression at either the transcriptional or posttranscriptional level. For example, a recent study showed that a *Salmonella* small noncoding RNA can discriminate mRNA regions that differ by a single nucleotide (Papenfort *et al.*, 2012). The picture is more complex for *ct058* and *ct192*. Regarding *ct058*, in L2/434 this gene is co-transcribed with *ct059*, and mRNA levels of *ct059* were also detected previously in ocular and urogenital strains (Borges *et al.*, 2010), which we confirmed (data not shown). We speculate that specific processing of the *ct059-ct058* transcript selectively reduces the levels of *ct058* mRNA in ocular and urogenital strains. Furthermore, *ct058* is a pseudogene in many ocular strains (but not in those used in the RT-qPCR assays) (Table 3.1 and also data not shown), and *ct192* is a pseudogene at least in the ocular strain B/Jali20 (Table 3.1). This does not necessarily have an impact on the interpretation of the data, as for example the LGV-specific pseudogene *ct300* showed similar mRNA levels during the developmental cycles of C/TW3, E/Bour, and L2/434 (Tables A.8, A.9, and A.10). Another issue is that there are different annotations in GenBank of the start codon of *ct192* in the archetypal ocular (A/Har13), urogenital (D/UW3), and LGV (L2/434) strains (Table 3.1). This alone could explain the differential expression of *ct192*, owing to dissimilar promoter regions. However, inspection of the nucleotide sequence immediately upstream from the annotated start codons of *ct192* in each of these three strains reveals a strong putative ribosome binding site only in L2/434 (Table 3.1). Therefore, it is likely that the start codon of *ct192* is conserved between *C. trachomatis* strains and corresponds to the one annotated in L2/434. In this situation, *ct192* is a pseudogene in all ocular strains (Table 3.1), and the differences in the mRNA levels of *ct192* between LGV and ocular or urogenital strains may be tentatively explained by the single-nucleotide differences found within its promoter region (Figure 3.7B and Figure A1B in Annexes).

It is not possible to make a rigorous side-by-side comparison between our RT-qPCR data and previous analyses of *inc* gene expression by RT-PCR (E. I. Shaw *et al.*, 2000) and microarrays (Belland *et al.*, 2003; Nicholson *et al.*, 2003), as the sensitivities of the methods used are quite different. In general, we have confirmed that *inc* genes display different profiles of expression but that they are mostly early-cycle genes, which supports the idea that many Inc proteins should play a role in modifying the inclusion membrane early during development (Valdivia, 2008; Betts *et al.*, 2009). We also found *inc* genes displaying an early- and late-cycle profile of expression. A similar profile has been observed in microarray studies (Belland *et al.*, 2003), and it was suggested to result from "carryover" mRNA of highly expressed late genes from the previous infectious cycle. However, the first time point we analyzed was at 2h postinfection, when carryover mRNA would have been degraded. Furthermore, at 30 or 42h postinfection, several late-cycle *inc* genes showed levels of expression at least comparable to those of early- and late-cycle *inc* genes (see Tables A.8, A.9, A.10, A.11, A.12, A.13, and A.14 in Annexes). This suggests that the expression of some *inc* genes could be induced early in the cycle, downregulated at mid-cycle, and induced again at late-cycle.

Although human macrophages have strong antimicrobial activity against *C. trachomatis* ocular and urogenital strains, they support the growth of LGV strains (Yong *et al.*, 1987). We hypothesize that specific amino acids in Inc proteins, or their earlier and/or higher expression, could specifically enable LGV strains to inhibit phagolysosomal fusion in macrophages and/or prevent the formation of reactive oxygen or nitrogen species (Flannagan *et al.*, 2012). Unfortunately, little is known about the function of the Inc proteins that we have identified as potentially involved in these processes. CT222 and CT118/IncG colocalize with kinases of the Src family in discrete regions of the

inclusion membrane that associate with host cell centrosomes (Mital *et al.*, 2010), CT223/IPAM and CT813/InaC reprogram the host microtubule network (Alzhanov *et al.*, 2009; Dumoux *et al.*, 2015; Kokes *et al.*, 2015), and CT228 might be involved in the mechanism of release of *C. trachomatis*. However, it is unclear how this may relate to our hypothesis. More recently, it was described a comprehensive Inc-human protein-protein interaction network (Mirrashidi *et al.*, 2015), where the authors show that CT058 might interact with elements of endosomes, the endoplasmic reticulum and mitochondria, and that CT192 might interact with elements of the centrosomes. Nevertheless, these interactions need further confirmation. On the other hand, Inc proteins that could manipulate intracellular membrane trafficking, such as those that have SNARE-like motifs (CT119/IncA) (Delevoye *et al.*, 2008) or interact with Rab GTPases (CT229) (Rzomp *et al.*, 2006), are good candidates to selectively inhibit macrophage phagolysosomal fusion. Recently, it was also shown that Incs can interact with each other (Mital *et al.*, 2013; Gauliard *et al.*, 2015). For example, CT222 was described to interact with himself, CT115/IncD, CT223, CT224, and CT850. Instead of a direct host target, Incs could form specific multi-protein complexes in the inclusion membrane, which in turn interact with host cell components or confer structure stability to the inclusion. Therefore, small variations in the amino acid sequence or in the gene expression of one Inc protein might not affect its function but the function of the entire complex. Not all *C. trachomatis* Inc proteins tentatively proposed to be specifically involved in inhibiting phagolysosomal fusion have homologues in other chlamydial species, which also avoid this host cell degradation pathway (Lutter *et al.*, 2010; Dehoux *et al.*, 2011). It is possible that each chlamydial species evolved particular Inc proteins, and other virulence proteins, to account for the specificity of each type of cell it infects.

3.6 Acknowledgements

This work was supported by Fundação para a Ciência e a Tecnologia (FCT) through grants PEst-OE/EQB/LA0004/2011 and PTDC/SAU-MII/099623/ 2008, and by the European Commission through a Marie Curie European Re-integration Grant (PERG03-GA-2008-230954). Filipe Almeida was a recipient of a PhD fellowship from FCT (SFRH/BD/73545/2010).

3.7 References

Albrecht, M., Sharma, C.M., Reinhardt, R., Vogel, J., and Rudel, T. (2009) Deep sequencing-based discovery of the *Chlamydia trachomatis* transcriptome. *Nucleic Acids Res* **38**: 868–877.

Belland, R.J., Zhong, G., Dean, D., Hogan, D., Sturdevant, D.E., Sharma, J., *et al.* (2003) Genomic transcriptional profiling of the developmental cycle of *Chlamydia trachomatis*. *Proc Natl Acad Sci U S A* **100**: 8478–8483.

Betts, H.J., Wolf, K., and Fields, K.A. (2009) Effector protein modulation of host cells: examples in the *Chlamydia* spp. arsenal. *Curr Opin Microbiol* **12**: 81–87.

Borges, V., Ferreira, R., Nunes, A., Nogueira, P.J., Borrego, M.J., and Gomes, J.P. (2010) Normalization strategies for real-time expression data in *Chlamydia trachomatis*. *J Microbiol Methods* **82**: 256–264.

Borges, V., Nunes, A., Ferreira, R., Borrego, M.J., and Gomes, J.P. (2012) Directional evolution of *Chlamydia trachomatis* towards niche-specific adaptation. *J Bacteriol* **194**: 6143–6153.

Brunelle, B.W., Nicholson, T.L., and Stephens, R.S. (2004) Microarray-based genomic surveying of gene polymorphisms in *Chlamydia trachomatis*. *Genome Biol* **5**: R42.

Caldwell, H.D., Wood, H., Dean, D., Bailey, R., Jones, R.B., Mabey, D., *et al.* (2003) Polymorphisms in *Chlamydia trachomatis* tryptophan synthase genes differentiate between genital and ocular isolates. *J Clin Invest* **111**: 1757–1769.

Carlson, J.H., Hughes, S., Hogan, D., Cieplak, G., Sturdevant, D.E., McClarty, G., *et al.* (2004) Polymorphisms in the *Chlamydia trachomatis* cytotoxin locus associated with ocular and genital isolates. *Infect Immun* **72**: 7063–7072.

Carlson, J.H., Porcella, S.F., Mcclarty, G., and Caldwell, H.D. (2005) Comparative Genomic Analysis of *Chlamydia trachomatis* Oculotropic and Genitotropic Strains. *Infect Immun* **73**: 6407–6418.

Dehoux, P., Flores, R., Dauga, C., Zhong, G., and Subtil, A. (2011) Multi-genome identification and characterization of *Chlamydiae*-specific type III secretion substrates: the Inc proteins. *BMC Genomics* **12**: 109.

Delevoye, C., Nilges, M., Dehoux, P., Paumet, F., Perrinet, S., Dautry-Varsat, A., and Subtil, A. (2008) SNARE protein mimicry by an intracellular bacterium. *PLoS Pathog* **4**.

Derré, I., Swiss, R., and Agaisse, H. (2011) The lipid transfer protein CERT interacts with the *Chlamydia* inclusion protein IncD and participates to ER-*Chlamydia* inclusion membrane contact sites. *PLoS Pathog* **7**.

Dumoux, M., Menny, A., Delacour, D., and Hayward, R.D. (2015) A *Chlamydia* effector recruits CEP170 to reprogram host microtubule organization. *J Cell Sci* **44**.

Fehlner-Gardiner, C., Roshick, C., Carlson, J.H., Hughes, S., Belland, R.J., Caldwell, H.D., and McClarty, G. (2002) Molecular basis defining human *Chlamydia trachomatis* tissue tropism: A possible role for tryptophan synthase. *J Biol Chem* **277**: 26893–26903.

Fields, K.A., Mead, D.J., Dooley, C.A., and Hackstadt, T. (2003) *Chlamydia trachomatis* type III secretion: Evidence for a functional apparatus during early-cycle development. *Mol Microbiol* **48**: 671–683.

Gauliard, E., Ouellette, S.P., Rueden, K.J., and Ladant, D. (2015) Characterization of interactions between inclusion membrane proteins from *Chlamydia trachomatis*. *Front Cell Infect Microbiol* **5**: 1–11.

Gomes, J.P., Nunes, A., Bruno, W.J., Borrego, M.J., Florindo, C., and Dean, D. (2006) Polymorphisms in the nine polymorphic membrane proteins of *Chlamydia trachomatis* across all serovars: Evidence for serovar da recombination and correlation with tissue tropism. *J Bacteriol* **188**: 275–286.

Hackstadt, T., Scidmore, M.A., Shaw, E.I., and Fischer, E.R. (1999) The *Chlamydia trachomatis* IncA protein is required for homotypic vesicle fusion. *Cell Microbiol* **1**: 119–130.

Harris, S.R., Clarke, I.N., Seth-Smith, H.M.B., Solomon, A.W., Cutcliffe, L.T., Marsh, P., *et al.* (2012) Whole-genome analysis of diverse *Chlamydia trachomatis* strains identifies phylogenetic relationships masked by current clinical typing. *Nat Genet* **44**: 413–419.

Jeffrey, B.M., Suchland, R.J., Quinn, K.L., Davidson, J.R., Stamm, W.E., and Rockey, D.D. (2010) Genome sequencing of recent clinical *Chlamydia trachomatis* strains identifies loci associated with tissue tropism and regions of apparent recombination. *Infect Immun* **78**: 2544–2553.

Joseph, S.J., Didelot, X., Gandhi, K., Dean, D., and Read, T.D. (2011) Interplay of recombination and selection in the genomes of *Chlamydia trachomatis*. *Biol Direct* **6**: 28.

Kokes, M., Dunn, J.D., Granek, J.A., Nguyen, B.D., Barker, J.R., Valdivia, R.H., and Bastidas, R.J. (2015) Integrating Chemical Mutagenesis and Whole-Genome Sequencing as a Platform for Forward and Reverse Genetic Analysis of *Chlamydia*. *Cell Host Microbe* 1–10.

Lutter, E.I., Barger, A.C., Nair, V., and Hackstadt, T. (2013) *Chlamydia trachomatis* Inclusion Membrane Protein CT228 Recruits Elements of the Myosin Phosphatase Pathway to Regulate Release Mechanisms. *Cell Rep* **3**: 1921–1931.

Lutter, E.I., Bonner, C., Holland, M.J., Suchland, R.J., Stamm, W.E., Jewett, T.J., *et al.* (2010) Phylogenetic analysis of *Chlamydia trachomatis* tarp and correlation with clinical phenotype. *Infect Immun* **78**: 3678–3688.

Lutter, E.I., Martens, C., and Hackstadt, T. (2012) Evolution and conservation of predicted inclusion membrane proteins in *Chlamydiae*. *Comp Funct Genomics* **2012**.

Mathews, S., and Timms, P. (2006) *In silico* identification of chlamydial promoters and their role in regulation of development. In *Chlamydia: genomics and pathogenesis*. Bavoil, P.M., and Wyrick, P.B. (eds). Horizon Bioscience, Norfolk, United Kingdom. pp. 133–156.

Mirrashidi, K.M., Elwell, C.A., Verschueren, E., Johnson, J.R., Frando, A., Dollen, J. Von, *et al.* (2015) Global Mapping of the Inc-Human Interactome Reveals that Retromer Restricts *Chlamydia* Infection. *Cell Host Microbe* **18**: 109–121.

Mital, J., Lutter, E.I., Barger, A.C., Dooley, C.A., and Hackstadt, T. (2015) *Chlamydia trachomatis* inclusion membrane protein CT850 interacts with the dynein light chain DYNLT1 (Tctex1). *Biochem Biophys Res Commun* **1**: 1–6.

Mital, J., Miller, N.J., Dorward, D.W., Dooley, C.A., and Hackstadt, T. (2013) Role for Chlamydial Inclusion Membrane Proteins in Inclusion Membrane Structure and Biogenesis. *PLoS One* **8**.

Nei, M., and Kumar, S. (2000) *Molecular evolution and phylogenetics*. Oxford University Press, New York, NY.

Nelson, D.E., Dean, D., Taylor, L.D., Dorward, D.W., Goheen, M.M., and Caldwell, H.D. (2006) Inhibition of *Chlamydiae* by primary alcohols correlates with the strain-specific complement of plasticity zone phospholipase D genes. *Infect Immun* **74**: 73–80.

Nicholson, T.L., Olinger, L., Chong, K., Schoolnik, G., and Stephens, R.S. (2003) Global stage-specific gene regulation during the developmental cycle of *Chlamydia trachomatis*. *J Bacteriol* **185**: 3179–3189.

Nunes, A., Gomes, J.P., Mead, S., Florindo, C., Correia, H., Borrego, M.J., and Dean, D. (2007) Comparative expression profiling of the *Chlamydia trachomatis pmp* gene family for clinical and reference strains. *PLoS One* **2**: 1–10.

Nunes, A., Nogueira, P.J., Borrego, M.J., and Gomes, J.P. (2008) *Chlamydia trachomatis* diversity viewed as a tissue-specific coevolutionary arms race. *Genome Biol* **9**: R153.

Papenfort, K., Podkaminski, D., Hinton, J.C.D., and Vogel, J. (2012) The ancestral SgrS RNA discriminates horizontally acquired *Salmonella* mRNAs through a single G-U wobble pair. *Proc Natl Acad Sci U S A* **109**: E757–E764.

Rzomp, K.A., Moorhead, A.R., and Scidmore, M.A. (2006) The GTPase Rab4 interacts with *Chlamydia trachomatis* inclusion membrane protein CT229. *Infect Immun* **74**: 5362–5373.

Saitou, N., and Nei, M. (1987) The Neighbor-joining Method: A New Method for Reconstructing Phylogenetic Trees. *Mol Biol Evol* **4**: 406–425.

Schachter, J. (1999) Infection and disease epidemiology. In *Chlamydia: intracellular biology, pathogenesis and immunity*. Stephens, R.S. (ed.). ASM Press, Washington DC. pp. 139–169.

Scidmore, M.A. (2005) Culture and Cultivation of *Chlamydia trachomatis*. *Curr Protoc Microbiol* 1–25.

Seth-Smith, H.M.B., Harris, S.R., Persson, K., Marsh, P., Barron, A., Bignell, A., *et al.* (2009) Co-evolution of genomes and plasmids within *Chlamydia trachomatis* and the emergence in Sweden of a new variant strain. *BMC Genomics* **10**: 239.

Shaw, A.C., Christiansen, G., Roepstorff, P., and Birkelund, S. (2000) Genetic differences in the *Chlamydia trachomatis* tryptophan synthase alfa-subunit can explain variations in serovar pathogenesis. *Microbes Infect* **2**: 581–592.

Shaw, E.I., Dooley, C.A., Fischer, E.R., Scidmore, M.A., Fields, K.A., and Hackstadt, T. (2000) Three temporal classes of gene expression during the *Chlamydia trachomatis* developmental cycle. *Mol Microbiol* **37**: 913–925.

Somboonna, N., Wan, R., Ojcius, D.M., Pettengill, M. a, Joseph, S.J., Chang, A., *et al.* (2011) Hypervirulent *Chlamydia trachomatis* Clinical Strain Is a Recombinant between Lymphogranuloma Venereum (L2) and D Lineages. *MBio* **2**: 1–12.

Stephens, R.S., Kalman, S., Lammel, C., Fan, J., Marathe, R., Aravind, L., *et al.* (1998) Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*. *Science* **282**: 754–759.

Subtil, A., Delevoye, C., Balañá, M.E., Tastevin, L., Perrinet, S., and Dautry-Varsat, A. (2005) A directed screen for chlamydial proteins secreted by a type III mechanism identifies a translocated protein and numerous other new candidates. *Mol Microbiol* **56**: 1636–1647.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**: 2731–2739.

Thomson, N.R., Holden, M.T.G., Carder, C., Lennard, N., Lockey, S.J., Marsh, P., *et al.* (2008) *Chlamydia trachomatis*: genome sequence analysis of lymphogranuloma

venereum isolates. *Genome Res* **18**: 161–171.

Unemo, M., Seth-Smith, H.M.B., Cutcliffe, L.T., Skilton, R.J., Barlow, D., Goulding, D., *et al.* (2010) The Swedish new variant of *Chlamydia trachomatis*: Genome sequence, morphology, cell tropism and phenotypic characterization. *Microbiology* **156**: 1394–1404.

Valdivia, R.H. (2008) *Chlamydia* effector proteins and new insights into chlamydial cellular microbiology. *Curr Opin Microbiol* **11**: 53–59.

CHAPTER 4 - *Chlamydia trachomatis* inclusion membrane protein CT288 interacts with the host cell centrosomal protein CCDC146

This Chapter contains data of a manuscript in preparation.
*Almeida, F, Pais, S.V., Mota, L.J. "*Chlamydia trachomatis *inclusion membrane protein CT288 interacts with the host cell centrosomal protein CCDC146".*

The author of this dissertation participated in all experiments described in this Chapter, with the exception of the generation of the* C. trachomatis *strain expressing CT288-2HA.

4.1 Abstract

After invasion of host cells, *Chlamydia trachomatis* resides within a membranaceous vacuolar compartment, known as the inclusion, and uses a type III secretion system to translocate several inclusion membrane (Inc) proteins that decorate the inclusion membrane. In Chapter 3, we hypothesized that a subgroup of Inc proteins might be determinant in the ability of some strains of *C. trachomatis* to infect macrophages and disseminate into lymph nodes. To test this hypothesis, we needed to further our understanding of the biological role of Inc proteins. For this, in this Chapter, we first used a yeast-two hybrid (Y2H) screen of a mammalian cDNA library to identify candidate interacting partners of *C. trachomatis* Incs CT228, CT249, and CT288. We further characterized the interaction between CT288 (563 amino acid residues) and the coiled coil domain containing protein 146 (CCDC146; 955 amino acid residues), a human protein of unknown function that localizes at the centrosome. By Y2H, we determined that the C-terminal regions of CT288 (amino acids 292-563) and of CCDC146 (amino acids 692-955) are necessary for the interaction. The interaction between CT288 ($CT288_{\Delta N \Delta TMD}$; deleted of its N-terminus and the hydrophobic motifs that characterize Inc proteins) and full-length of CCDC146 or $CCDC146_{692\text{-}955}$ was further validated by immunoprecipitation experiments after ectopic expression of the proteins in mammalian cells. Furthermore, ectopically expressed $CCDC146_{692\text{-}955}$ can interact with full-length CT288 expressed by *C. trachomatis* during infection. Immunofluorescence microscopy of mammalian cells infected by *C. trachomatis* showed that ectopically expressed CCDC146 localizes at the inclusion membrane. Thus, our data reveals a new host cell target for an Inc protein, possibly involved in the previous described interactions between the inclusion and the host centrosome.

4.2 Introduction

Throughout the developmental cycle *Chlamydia trachomatis* maintains an intact membranaceous compartment (inclusion) and manipulates host cellular processes via the delivery of type III secretion (T3S) effector proteins directly into the host cytosol and membranes (Valdivia, 2008; Betts *et al.*, 2009). Among these effectors, the Inc proteins have a privileged localization at the inclusion membrane, likely directed by a characteristic bi-lobed hydrophobic domain (Li *et al.*, 2008; Mital *et al.*, 2010; Dehoux *et al.*, 2011).

Host cell interacting partners have been found for several *C. trachomatis* Inc proteins, which revealed possible effector functions. For example, CT228 interacts with the myosin phosphatase target subunit 1 (MYPT1), an element of the myosin phosphatase pathway, favoring the extrusion mechanism for the bacterial egress from the host cell (Lutter *et al.*, 2013); CT229 binds Rab4, suggesting that it might play a role in regulating the intracellular trafficking or fusogenicity of the chlamydial inclusion (Rzomp *et al.*, 2006); CT115/IncD interacts with the ceramide transfer protein (CERT), which is involved in the non-vesicular transfer of ceramide from the endoplasmic reticulum (ER) to the Golgi (Derré *et al.*, 2011; Agaisse and Derré, 2014); CT119/IncA binds to VAMP3 and VAMP8 through a SNARE domain and promotes fusion of membranes during infection (Hackstadt *et al.*, 1999; Delevoye *et al.*, 2008); CT118/IncG interacts with the phosphoserine-binding protein 14-3-3-β, which is involved in host signal-transduction pathways (Scidmore and Hackstadt, 2001); CT850 binds dynein light chain (DYNLT1) and promotes the positioning of the inclusion at the centrosomal region (Mital *et al.*, 2015); CT223/IPAM (inclusion protein acting on microtubules) and CT813/InaC (inclusion membrane protein for actin assembly) interact with the host centrosomal protein CEP170 and ARF1, respectively, and actively participate in the modification of the microtubule organization

around the inclusion (Dumoux *et al.*, 2015; Kokes *et al.*, 2015); CT116/IncE interacts with sorting nexins, leading to the disruption of the retromer function (Mirrashidi *et al.*, 2015). In this last study, Mirrashidi and colleagues used 58 purified Inc proteins to affinity purify host cell interacting partners followed by their identification by mass spectrometry (Mirrashidi *et al.*, 2015). This revealed a comprehensive Inc-human protein-protein interaction network, where published Inc-host protein interactions have been validated and high-confidence host cell targets were identified for 38 Inc proteins (Mirrashidi *et al.*, 2015). In summary, these discoveries further indicated that Inc proteins are important factors to mediate the interaction between the inclusion and the host cell.

We previously identified the majority of Incs as T3S substrates (Chapter 2), and hypothesized that a subset of Inc proteins could contribute to the unique tropism and invasiveness of *C. trachomatis* lymphogranuloma venereum (LGV) strains (serovars L1-L3), relative to ocular (serovars A-C) and urogenital (serovars D-K) strains (Chapter 3). To test this hypothesis, in this Chapter, we aimed to unveil the biological role of Incs from this subset of proteins (Table 3.3 in Chapter 3), by performing yeast-two hybrid (Y2H) screens of a mammalian cDNA library to search for putative interacting partners of Incs CT228, CT249, and CT288, for which host cell interacting partners were unknown at the time. We found an interaction between Inc CT288 and the human centrosomal protein CCDC146, which was further validated biochemically by co-immunoprecipitation experiments. We also showed by immunofluorescence microscopy that both proteins co-localize at the inclusion membrane during infection by *C. trachomatis*.

4.3 Materials and Methods

4.3.1 Mammalian cell lines

HeLa 229 cells (from the ATCC) were used to propagate *C. trachomatis*, and for infections with *C. trachomatis*. HEK293T cells (from the ATCC) were used for the experiments involving transfections for co-IPs. Both cell lines were maintained at 37°C in a 5% CO_2 atmosphere in DMEM (Life Technologies) supplemented with 10% (v/v) FBS (Life Technologies). When HEK293T cells were seeded for co-IP experiments, the tissue culture plates were previously coated with 0.001% (v/v) poly-L-lysine (Sigma-Aldrich) in PBS.

4.3.2 Bacterial strains and growth conditions

Escherichia coli TOP10 (Life Technologies) was used for standard cloning methods and for plasmid amplification, and the methylation deficient *E. coli* K12 ER2925 (Dam and Dcm negative; New England Biolabs) was used to purify plasmids for *C. trachomatis* transformation. *E. coli* strains were routinely grown at 37ºC in LB medium with the appropriate antibiotics and supplements. Plasmids were introduced into *E. coli* TOP10 and into *E. coli* K12 ER2925 by standard methods (electroporation or $CaCl_2$ transformation).

4.3.3 Plasmids and DNA primers

The plasmids used and constructed in this Chapter are listed in Table A.2, and primers are listed in Table A.1 (Annexes). The construction of the plasmids is detailed in Table A.2 (Annexes). Plasmids were constructed and purified using proofreading Phusion DNA polymerase, restriction enzymes, T4 DNA Ligase, and DreamTaq DNA polymerase (all from Life Technologies), a DNA Clean & Concentrator-5 (Zymo

Research), a Zymoclean Gel DNA Recovery Kit (Zymo Research), and a GeneElute plasmid miniprep kit (Sigma). The accuracy of the nucleotide sequence of the inserts in the constructed plasmids was verified by DNA sequencing (Stab Vida).

4.3.4 Y2H screens

The Matchmaker® Gold Yeast Two-Hybrid System (Clontech) was used for the Y2H screen of a pre-transformed Mate & Plate™ Library – normalized Universal Human HeLa cDNA library (Clontech), following the instructions of the manufacturer. In summary, the yeast *Saccharomyces cerevisiae* Y2HGold was transformed with plasmid derivatives from pGBKT7 (Table A.2 in Annexes) and mated with yeast strain Y187 carrying the HeLa cDNA library cloned into pGADT7. The two strains were mated and plated in lower stringency or double dropout (DDO) media (SD/–Leu/–Trp) supplemented with X-α-Galactosidase (X) and Aureobasidin A (A) (all from Clontech), and incubated 3-5 days at 28-30ºC. The blue colonies grown on DDO/X/A media were patched into a higher stringency or quadruple dropout (QDO) media (SD/–Ade/–His/–Leu/–Trp) supplemented with X-α-Galactosidase and Aureobasidin A, and incubated 3-5 days at 28-30ºC. The blue colonies that grew in QDO/X/A were further analyzed for autonomous system activation, and for identification of the target protein, by sequencing the plasmid DNA (at Stab Vida) with primers listed in Table A-1 (Annexes). For CT228, we analyzed $5.87x10^5$ clones, patched 1 blue colony and recovered 1 from high stringent media for further analysis. For CT249, we analyzed $5.78x10^6$ clones, patched 60 blue colonies and recovered 19 from high stringent media for further analysis. For CT288, we analyzed $1.43x10^8$ clones, patched 60 blue colonies and recovered 41 from high stringent media for further analysis.

4.3.5 Preparation of yeasts extracts

Protein extracts of the *S. cerevisiae* Y2HGold strains expressing Gal4BD fusion proteins were prepared by the TCA method (Matchmaker® Gold Yeast Two-Hybrid System supporting protocols). Briefly, an overnight culture was incubated at 28-30°C with agitation (130 rpm) for 6 h. The final OD_{600} was measured and the cells were chilled on ice and centrifuged at 1000 *g* for 5 min and at 4ºC. The cell pellets were washed with ice-cold sterile H_2O (MilliQ®), centrifuged again, and immediately stored at -80ºC. Then, each cell pellet was resuspended in 100 µl ice-cold TCA buffer per 7.5 OD_{600} unit (20 mM Tris pH 8.0, 50 mM ammonium acetate, 2 mM EDTA, and freshly added 0.1 mM PMSF), 100 µl ice-cold 20% TCA (v/v) per 7.5 OD_{600} unit and approximately 100 µl of 500 µm glass beads (Sigma). To disrupt the yeast cell wall, the cell pellets were subjected to vortex for 4 periods of 1 min each and at room temperature, placing the tubes on ice for 30 secs in between each vortexing step. The supernatant above the settled glass beads was transferred to a fresh 1.5 ml tube and the beads were washed in a 1:1 mixture of TCA buffer and 20% TCA. The tubes with the washed beads were vortexed again for 2 times during 1 min and at room temperature. The supernatant above the glass beads was collected and joined with the first supernatant. The proteins in the supernatant were pelleted by centrifugation at 17000 *g* for 10 min and at 4°C. The pellets were resuspended in 10 µl of TCA-Laemmli buffer per OD_{600} unit (to prepare 1 ml, added 480 µl SDS/glycerol solution [SDS 7.3% (w/v), glycerol 29.1% (v/v), Tris-base 83.3 mM, a spatula tip-full of bromophenol blue], 400 µl Tris/EDTA [200 mM Tris-base and 20 mM EDTA], 50 µl β-mercaptoethanol, 20 µl of 10 mM PMSF, and 50 µl MilliQ® H_2O). Denatured the samples 10 min at 100 ºC, and centrifuged 10 min at 17000 *g,* and at room temperature. Recovered the supernatant and proceeded to immunoblotting.

4.3.6 Manipulation of C. trachomatis

C. trachomatis prototype serovar L2 strain 434/Bu (L2/434; from the ATCC) was maintained and propagated in HeLa 229 cells using standard techniques (Scidmore, 2005). For immunofluorescence microscopy (see below), cells were seeded at $5x10^4$ cells per well on glass coverslips, and for immunoblotting of whole cell extracts (see below), cells were seeded at $1x10^5$ cells per well, in all cases in 24-well tissue culture plates. HeLa 229 cells were equilibrated in HBSS and infected or mock-infected at a MOI of 1. After an incubation of 30 min at 37ºC in a 5% CO_2 atmosphere, the inoculum was removed and fresh DMEM containing 50 µg/ml of gentamicin was added. At this point and where appropriate, cells were transfected with plasmids encoding for EGFP, EGFP-CCDC146$_{1-955}$, or CCDC146$_{1-955}$-HA (Table A.2 in Annexes). Twenty-four hours post-infection, the cells were collected for whole cell extracts for immunoblotting, or fixed with 1% (v/v) PFA or methanol for immunofluorescence microscopy. For the co-IP experiments, HEK293T cells were seeded at $5x10^6$ per well of a 6-well tissue culture plate. The cells were infected with *C. trachomatis* transformed with pSVP255 (see below) at a MOI of 2, and transfected with plasmids encoding for EGFP, EGFP-CCDC146$_{692-955}$, or EGFP-CCDC146$_{1-955}$ as explained above for infection of Hela 229 cells. For the transformation of *C. trachomatis*, Hela 229 cellular extracts containing *C. trachomatis* elementary bodies (EBs) were isolated by sonication from a 44 h infection with the L2/434 strain. The lysate was centrifuged and the supernatant was frozen at -80°C in 4SP (sucrose phosphate buffer) 0.4 mM sucrose, 17 mM Na_2HPO_4, 3 mM NaH_2PO_4 and pH7.4). For titrations, HeLa 229 cells were seeded onto glass coverslips and infected with several dilutions of the cellular extracts containing EBs. At 30 h post-infection, cells were fixed, immunolabeled, and inclusions were enumerated by immunofluorescence microscopy.

4.3.7 C. trachomatis *transformation*

C. trachomatis L2/434 was transformed with plasmid pSVP255 (Table A.2 in Annexes), encoding full-length CT288 with a C-terminal 2HA tag ($CT288_{1-563}$-2HA). Transformation of *C. trachomatis* was performed essentially as previously described (Agaisse and Derré, 2013), but with some modifications. First, 20 µl of a cellular extract containing *C. trachomatis* EBs stored in 4SP were added to 6 µg of plasmid DNA diluted in 200 µl of $CaCl_2$ buffer (10 mM Tris pH 7.4, 50 mM $CaCl_2$). The mixture was then homogenized using a vortex, and incubated for 30 min at room temperature. Meanwhile, $4x10^6$ HeLa 229 cells were trypsinized and centrifuged for 5 min at 237 *g* and at room temperature. Cells were washed with PBS and resuspended in 200 µl of $CaCl_2$ buffer. The cells were added to the mixture and incubated 20 min at room temperature. One hundred 100 µl of the whole mixture (cells, plasmid DNA, and *Chlamydia*) were transferred to 3 ml of pre-warmed DMEM in a 8.5 cm^2 Petri dish and incubated for 44 h at 37ºC in a 5% CO_2 atmosphere. The medium was then removed and the cells were osmotically lysed with 2 ml of sterile H_2O (MilliQ®). The lysate was centrifuged for 5 min at 237 *g*, and at room temperature, after which 2 ml of the supernatant were added to 2 ml of sucrose-phosphate-glutamate buffer (SPG; 0.2 mM sucrose, 17 mM Na_2HPO_4, 3 mM NaH_2PO_4, 5 mM L-glutamic acid). This was then used as inoculum to add to newly seeded HeLa 229 cells ($4x10^6$ cells in a 75 cm^2 surface area flask), previously equilibrated in HBSS. After an incubation of 1h at room temperature with gentle rocking, the inoculum was removed, and DMEM supplemented with 0.3 U/ml penicillin and 1 µg/ml cycloheximide was added. The cells were then incubated for 44 h at 37ºC in a 5% CO_2 atmosphere. The infected cells were collected by scrapping in 3 ml SPG, and lysed by sonication (2 pulses of 20 s with 100 W and a cycle of 0.5). The lysate was then centrifuged for 5 min at 237 *g* and at room temperature, after which, the supernatant was split in two and used to re-infect newly seeded HeLa

229 cells ($1.3x10^6$ cells in two 25 cm^2 surface area flasks), previously equilibrated in HBSS. After an incubation of 1 h at room temperature and gentle rocking, the inoculum was removed, and DMEM supplemented with 0.3 U/ml penicillin and 1 µg/ml cycloheximide was added. The cells were then incubated for 3 days at 37ºC in a 5% CO_2 atmosphere. The infected cells in both flasks were collected by scrapping in 1.5 ml SPG, and lysed by sonication followed by centrifugation as described above. The whole supernatant was used to re-infect newly seeded HeLa 229 cells ($1.3x10^6$ cells in a 25 cm^2 surface area flask), previously equilibrated in HBSS. After an incubation of 1 h at room temperature, the inoculum was removed, and DMEM supplemented with 1 U/ml penicillin and 1 µg/ml cycloheximide was added. The cells were incubated for 3 days at 37ºC in a 5% CO_2 atmosphere. At this point, the presence of wild type inclusions, as detected by phase-contrast microscopy, indicated a successful transformation. The infected cells containing transformed *C. trachomatis* were collected by scrapping in 1.5 ml SPG, and lysed by sonication followed by centrifugation as described above, after which an adequate amount of the supernatant (depending on the outcome of the previous infection) was used to re-infect newly seeded HeLa 229 cells ($1.3x10^6$ cells in a 25 cm^2 surface area flask), previously equilibrated in HBSS. After an incubation of 1 h at RT, the inoculum was removed, and DMEM supplemented with 1 U/mL penicillin and 1 µg/mL cycloheximide was added. The cells were incubated 44 h at 37ºC in a 5% CO_2 atmosphere. The transformed *C. trachomatis* was continually passaged for approximately 10 rounds of re-infection, then stored at -80ºC and titrated for further experiments.

4.3.8 Transient transfection of mammalian cells

HeLa 229 and HEK293T cells were transfected with plasmid DNA by using the jetPEI reagent (Polyplus-Transfection) as detailed in the

instructions, but using 250 ng of DNA per well of a 24-well tissue culture plate or 1250 ng of DNA per well of a 6-well tissue culture plate. Cells were transfected immediately after infection with *C. trachomatis* and incubated at 37°C in a 5% CO_2 atmosphere for the indicated time periods prior to collection and immunoblotting, or fixation and immunolabeling.

4.3.9 Preparation of cell lysates, fractionation and co-IP

For the co-IP, we used the GFP Trap kit (Chromotek), according to the protocol of the manufacturer with minor adjustments. In experiments involving only co-transfection, $5x10^6$ HEK293T cells per well of a 6-well tissue culture plate, in a total of 6 wells per condition, were transfected with combinations of two plasmids, one encoding for EGFP fusion proteins and the other encoding for HA-tagged proteins. In experiments involving infection and transfection, $5x10^6$ HEK293T cells per well of a 6-well tissue culture plate, in a total of 12 wells per condition, were infected with *C. trachomatis* transformed with pSVP255 at a MOI of 2, and transfected with plasmids encoding for EGFP, EGFP-CCDC146$_{692-955,}$ or EGFP-CCDC146$_{1-955}$ (Table A.2 in Annexes). In both cases, the cells were collected by trypsinization after a period of incubation of 24 h at 37ºC in a 5% CO_2 atmosphere. The cells were pelleted by centrifugation and washed with PBS. In the case of the infection and transfection experiments, an additional step of cross-link with paraformaldehyde was done to increase the possibility of capturing the interaction between CT288 and CCDC146. For that, the cell pellets were resuspended in 1 ml of paraformaldehyde 1% (w/v) in PBS per $1x10^7$ cells and incubated 7 min at room temperature. The cells were immediately centrifuged for 3 min at 1800 *g* and at room temperature. Then, washed the pellets two times with 500 µl 1.25 M glycine, and centrifugations of 3 min at 1800 *g* and at room temperature. The remainder of the protocol was identical for

co-transfection, and infection and transfection experiments. The cell pellets were lysed in 200 µl ice-cold co-IP Lysis buffer (20 mM Tris-HCl pH 7.5, 137 mM NaCl, 2 mM EDTA, 1.0% NP40, freshly added protease Inhibitors and PMSF) for 30 min on ice, and mixed by pipetting every 10 min. The lysates were centrifuged for 10 min at 17000 *g* and at 4ºC, and the supernatants were either used for analysis in SDS-PAGE (input of co-IP) or added to 800 µl of ice-cold GFP Trap buffer (10 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA). Then, 750 µl of the diluted supernatant were added to previously washed GFP Trap beads and incubated overnight at 4ºC with end-over-end mixing. Washed the beads six times with GFP Trap buffer and centrifugations of 2 min at 2000 *g* and at 4ºC. Finally, the pelleted beads (output of co-IP) and the input of co-IP were resuspended in 20 µl SDS-PAGE Laemmli Buffer 2X [Tris-HCl 100 mM, pH 6.8, SDS 4.0% (w/v), glycerol 20% (v/v), β-mercaptoethanol 0.2 M, bromophenol blue 0.2% (w/v)]. Denatured the samples 10 min at 100ºC and proceeded to immunoblotting.

4.3.10 Antibodies

For immunoblotting, mouse anti-myc antibody (Calbiochem) was used at 1/1000, mouse anti-α-tubulin (Sigma-Aldrich) was used at 1/1000, rat anti-HA 3F10 (Roche) was used at 1/1000, goat anti-GFP (Sicgen) was used at 1/1000, and goat anti-MOMP (Abcam) was used at 1/1000. Horseradish peroxidase-conjugated secondary antibody anti-mouse (GE Healthcare) was used at 1/10000, anti-rat (Sigma-Aldrich) was used at 1/10000, and anti-goat (Jackson ImmunoResearch Laboratories) was used at 1/10000. For immunofluorescence, mouse anti-γ-tubulin antibody (Sigma-Aldrich) was used at 1:200, mouse anti-CT442 antibody (Li *et al.*, 2008) was used at 1:200, rat anti-HA 3F10 antibody (Roche) was used at 1:200, goat anti-GFP antibody (Sicgen) was used at 1:200, and goat anti-MOMP antibody (Abcam) was used at 1:200. Secondary

antibody donkey anti-goat conjugated to cyanine 5 (Jackson ImmunoResearch Laboratories) was used at 1:200, anti-rat conjugated to rhodamine RedX (Jackson ImmunoResearch Laboratories) was used at 1:200, and goat anti-mouse AF568 (Invitrogen) was used at 1:200.

4.3.11 Immunoblotting

To prepare total mammalian cell extracts, HeLa 229 cells were trypsinized, collected and centrifuged 5 min at 2300 *g* and at 4ºC. The pellet was washed twice with ice-cold PBS, and resuspended in 50 µl SDS-PAGE Laemmli Buffer 1X [Tris-HCl 50 mM, pH 6.8, SDS 2.0% (w/v), glycerol 10% (v/v), β-mercaptoethanol 0.1 M, bromophenol blue 0.1% (w/v)] and 1 µl benzonase. The proteins were denatured 10 min at 100ºC. Total yeast cell extracts, total mammalian cell extracts, and fractionated co-IP samples were resolved in 12% SDS-PAGE, followed by immunoblotting. Proteins in the gel were transferred onto PVDF membranes (Bio-Rad) and blocked in 5% (w/v) dried skimmed milk diluted in PBS containing 0.1% (v/v) Tween-20. The membranes were probed with primary and horseradish peroxidase-conjugated secondary antibodies, and detected using the Western Lightning Plus-ECL kit (Perkin Elmer) in a Chemidoc XRS+ system (Bio-Rad) or by exposure to Amersham Hyperfilm ECL (GE Healthcare) in a dark room.

4.3.12 Immunofluorescence microscopy

For immunofluorescence microscopy, cells were either fixed with paraformaldehyde 4% (w/v) 15 min at RT or with methanol 5 min at -20ºC (whenever the antibody for γ-tubulin was used). Antibodies were diluted in PBS with 0.1% (v/v) Triton-X100 and 10% (v/v) horse serum. The cells in the coverslips were washed in PBS with 0.1% (v/v) Triton-X100 and incubated for 1 h with primary antibodies. The cells in the

coverslips were washed again in PBS with 0.1% (v/v) Triton-X100 and incubated with appropriate secondary antibodies for another 1 h. The cells in the coverslips were washed again first in PBS with 0.1% (v/v) Triton-X100, then PBS, and finally in sterile H_2O (MilliQ®). In the end, coverslips were mounted onto glass slides using Aqua-poly/Mount mounting medium (Polysciences). Samples were analyzed using a widefield fluorescence microscope (Leica DMRA2) or a confocal laser scanning microscope (Zeiss LSM 710 META) at the Instituto Gulbenkian da Ciência (IGC) or at the Centro de Doenças Crónicas (CEDOC), respectively. All images were obtained by confocal microscopy and processed using Zeiss LSM Image Browser and Adobe Photoshop software.

4.4 Results

4.4.1 Y2H screen using Inc proteins as baits for a mammalian cDNA library

To identify host cell interacting partners of Inc proteins, we performed Y2H screens using Inc proteins CT228, CT249, and CT288 as bait for fragments of proteins expressed by a mammalian cDNA library. As Inc proteins possess bi-lobed hydrophobic regions, we used plasmids encoding the DNA-binding domain of the yeast transcription factor Gal4 containing a myc epitope at the C-terminal (Gal4BD-myc) fused to CT228 (amino acids 1-591), CT249 (amino acids 1-351), or CT288 (amino acids 1-563) deleted of their predicted transmembrane domains (TMDs) (Dehoux *et al.*, 2011). In particular, for CT228 and CT249, we constructed plasmids encoding CT228 without amino acids 38 to 86, and CT249 without amino acids 51 to 97 (Gal4BD-myc-CT228$_{\Delta TMD}$ and Gal4BD-myc-CT249$_{\Delta TMD}$ proteins; Figure 4.1A). For CT288, we constructed a plasmid encoding CT288 lacking the first 35 amino acids and the two bi-lobed hydrophobic domains (amino acids 36 to 88 and 242 to 291; Gal4BD-myc-CT288$_{\Delta N\Delta TMD}$; Figure 4.1A). Each plasmid was introduced in *S. cerevisiae* Y2HGold and expression of the Gal4BD-myc-Inc proteins was analyzed by immunoblotting of yeast protein extracts using an anti-myc antibody. Figure 4.1B shows that all fusion proteins were expressed and migrated on SDS-PAGE as expected from their predicted molecular mass of 32 kDa (Gal4BD-myc-CT228$_{\Delta TMD}$), 23 kDa (Gal4BD-myc-CT249$_{\Delta TMD}$), and 66 kDa (Gal4BD-myc-CT288$_{\Delta N\Delta TMD}$). We also verified that the Gal4BD-myc-Inc proteins did not autonomously activate the reporter genes in the absence of a prey protein, by plating the transformed Y2HGold strains in normal and in higher stringency medium (Figure 4.2A, B and C; upper panels). Based on these results, we selected the plasmids encoding the CT228$_{\Delta TMD}$, CT249$_{\Delta TMD}$, and CT288$_{\Delta N\Delta TMD}$ Gal4BD-myc fusion proteins for the Y2H screens.

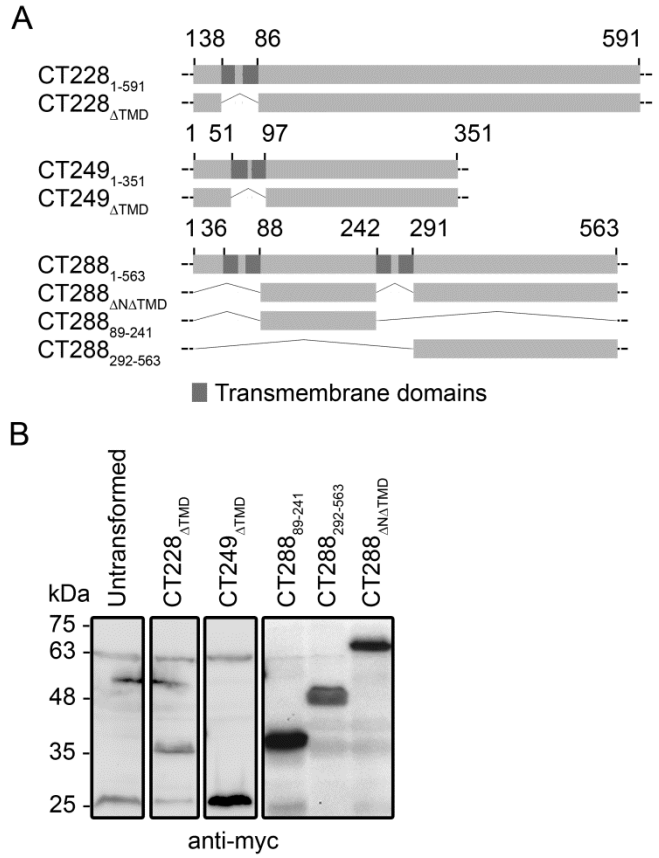


Figure 4.1 – Fragments of the *C. trachomatis* Inc proteins CT228, CT249 and CT288 used in the Y2H screen. (A) Domain organization of CT228, CT249, and CT288 fused to the Gal4 DNA-binding domain containing a myc epitope at the C-terminal (Gal4BD-myc-Inc). (B) Immunoblot of yeast protein extracts (see materials and methods) from $3x10^7$ yeast cells (considering that 1 OD_{600} = $3x10^7$ yeast cells) after expression of CT228, CT249, and CT288 Gal4BD-myc-Inc fusion proteins.

The yeast strains harboring each plasmid were used as bait to screen the yeast strain Y187 containing a normalized universal human HeLa cDNA library constructed in pGADT7 as fusions to the Gal4 activation domain containing a HA epitope at the C-terminal (Gal4AD-HA-prey). The interacting prey plasmids were isolated and tested for autonomous system activation. This showed that the only interaction found for CT228 was not a false positive, that only 4 of the 19 interactions found for

CT249 were not false positives, and that 38 of the 41 interactions found for CT288 were not false positives (examples in Figure 4.2). Next, among the real positive interactions, we looked for in-frame fusions to the GAL4AD-HA by sequencing the inserts in the plasmids. A BLAST analysis of the DNA sequences revealed in-frame fusions for the only interaction found for CT228, for 3 of the 4 real interactions found for CT249, and for 28 of the 38 real interactions found for CT288 (Table 4.1).

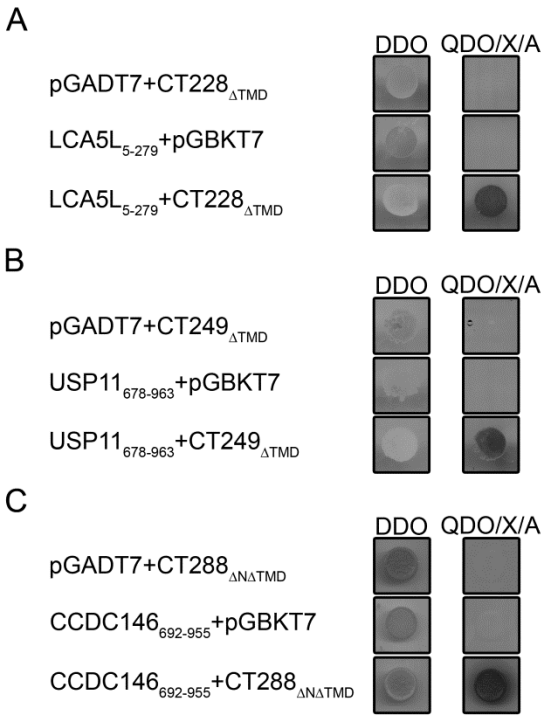


Figure 4.2 – Yeast two-hybrid screen of a mammalian cDNA library using portions of CT228, CT249, and CT288. Examples of the interactions found in the Y2H screen using the Incs CT228 (A), CT249 (B), and CT288 (C). The interactions were verified by plating the yeast strain Y2HGold harboring different combinations of plasmids encoding Gal4BD-myc-Inc fusions of CT228, CT249, and CT288, the Gal4AD-HA-prey fusion of the putative interacting partners, or empty plasmids in normal DDO medium, and in higher stringency QDO/X/A medium. LCA5L is the Leber-congenital amaurosis 5 like protein, USP11 is the ubiquitin specific peptidase 11, and CCDC146 is the human coiled coil domain containing protein 146 (Table 4.1).

Table 4.1– Candidate binding partners of CT228, CT249, and CT288 identified in the Y2H screen of a mammalian cDNA library.

Inc Protein	Y2H candidate interacting partners	In-frame fusion (amino acids)	Number of Hits
CT228	LCA5L – Leber-congenital amaurosis 5 like protein	Yes (5-279)	1 in 1
CT249	USP11 – Ubiquitin specific peptidase 11	Yes (678-963)	1 in 4
	PEX1 – Peroxisomal biogenesis factor 1	Yes (897-1283)	1 in 4
	NOL11 – Nucleolar protein 11	Yes (302-719)	1 in 4
	Homo sapiens 3 BAC RP11-65D10	No	1 in 4
CT288	CCDC146 – Coiled coil domain containing protein 146	Yes (692-955)	9 in 38
	LCA5L – Leber-congenital amaurosis 5 like protein	Yes (5-279)	2 in 38
	Homo sapiens anterior gradient 3 homolog (*Xenopus laevis*) (AGR3)	Yes (1-166)	1 in 38
	Homo sapiens coiled-coil domain containing 42 (CCDC42)	Yes (6-242)	1 in 38
	Homo sapiens inhibitor of kappa light polypeptide gene enhancer in B-cells (IKBKAP)	Yes (1156-1331)	1 in 38
	Homo sapiens fasciculation and elongation protein zeta 2 (zygin II) (FEZ2)	Yes (181-353)	1 in 38
	Homo sapiens THO complex 7 homolog (Drosophila) (THOC7)	Yes (1-254)	1 in 38
	Homo sapiens exocyst complex component 1 (EXOC1)	Yes (120-358)	1 in 38
	Homo sapiens MLF1 interacting protein (CENP-50)	Yes (194-418)	1 in 38
	Homo sapiens catenin (cadherin-associated protein)	Yes (690-860)	1 in 38
	Homo sapiens COP9 constitutive photomorphogenic homolog subunit 4	Yes (1-406)	1 in 38
	Homo sapiens protein phosphatase 2, regulatory subunit B' (PP2A)	Yes (274-429)	1 in 38
	Homo sapiens structural maintenance of chromosomes 3 (SMC3)	Yes (828-1217)	1 in 38
	Homo sapiens superkiller viralicidic activity 2-like	Yes (895-1042)	1 in 38
	Homo sapiens interferon-induced protein 44-like (IFI44L)	Yes (327-452)	1 in 38
	Homo sapiens mediator complex subunit 4	Yes (1-270)	1 in 38
	Homo sapiens fasciculation and elongation protein zeta 1 (zygin I)	Yes (238-392)	1 in 38
	Homo sapiens inhibitor of growth family, member 5 (ING5)	Yes (1-130)	1 in 38
	Homo sapiens nucleoporin 88kDa (NUP88)	Yes (517-741)	1 in 38
	Homo sapiens furry homolog (Drosophila) (FRY)	Yes (2782-3013)	1 in 38

Inc Protein	Y2H candidate interacting partners	In-frame fusion (amino acids)	Number of Hits
CT288	Homo sapiens translocated promoter region, nuclear basket protein (TPR)	No	2 in 38
	Homo sapiens heparan sulfate 6-O-sulfotransferase 2 (HS6ST2)	No	1 in 38
	Homo sapiens receptor accessory protein 5	No	1 in 38
	Homo sapiens tripartite motif containing 22 (TRIM22)	No	1 in 38
	Homo sapiens spermatogenic leucine zipper 1	No	1 in 38
	Homo sapiens solute carrier family 35	No	1 in 38
	Homo sapiens bromodomain adjacent to zinc finger domain	No	1 in 38
	DNA region of the chromosome 8, nothing specific	No	1 in 38

As the human coiled-coil domain containing protein 146 (CCDC146) appeared in a high number of hits in the Y2H screen using CT288 as bait (9 of the 28 in-frame hits), we decided to characterize in more detail the putative interaction between CT288 and CCDC146.

CCDC146 is a 955 amino acids protein of unknown function, and the region encoded in the recovered prey plasmids contained its C-terminal region (amino acids 692 to 955) (Figure 4.2C). A BLAST analysis with this protein did not reveal significant identity to known protein domains, with the exception of a structural maintenance of chromosomes (SMC) domain (Strunnikov and Jessberger, 1999). CCDC146 was recently described to be localized at the centrosome of mammalian cells (Firat-Karalar *et al.*, 2014).

4.4.2 Characterization of the interaction between CT288 and CCDC146 by Y2H

We next used Y2H to identify the region of CT288 involved in binding to the C-terminal region of CCDC146 (amino acids 692 to 955). For that, we constructed plasmids encoding the Gal4BD-myc fused to amino

acids 89-241 of the N-terminal and to amino acids 292-563 of the C-terminal region of CT288 (Gal4BD-myc-CT288$_{89-241}$ and Gal4BD-myc-CT288$_{292-563}$, respectively) (Figure 4.1A). We confirmed that the Gal4BD-myc-Inc fusion proteins were expressed and migrated on SDS-PAGE according to their predicted molecular mass of 34 and 47 kDa (Gal4BD-myc-CT288$_{89-241}$ and Gal4BD-myc-CT288$_{292-563}$, respectively) (Figure 4.1B). We also confirmed that the Gal4BD-myc-Inc proteins did not autonomously activate the system (Figure 4.3; upper panels).

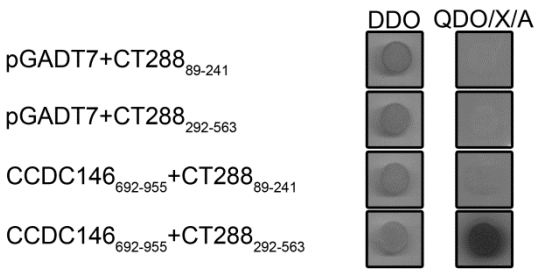


Figure 4.3 – The C-terminal region of CT288 is responsible for the interaction with the C-terminal region of CCDC146. The **i**nteraction between the two regions of CT288 and the C-terminal region of CCDC146 were verified by plating the yeast strain Y2HGold harboring different combinations of plasmids encoding the Gal4BD-myc fusions of CT288, the Gal4AD-HA fusion of CCDC146 (Gal4AD-HA-CCDC146$_{692-955}$), or empty plasmids in normal DDO medium, and in higher stringency QDO/X/A medium.

Combinations of two plasmids, one encoding for the Gal4BD-myc fusions of CT288 or empty plasmid (pGBKT7), and the other encoding for the Gal4AD-HA CCDC146 fusion (Gal4AD-HA-CCDC146$_{692-955}$) or empty plasmid (pGADT7) were introduced into *S. cerevisiae* Y2HGold. Then, the yeast strains were plated, in parallel, in normal and in higher stringency medium. We only observed growth in form of blue colonies when both CT288$_{292-563}$ and CCDC146$_{692-955}$ were present, indicating that the region of CT288 responsible for the binding to CCDC146 lies between amino acids 292 and 563 (Figure 4.3; bottom panels).

4.4.3 CT288 from C. trachomatis *strains L2/434, C/TW3, and E/Bour interact with CCDC146 by Y2H*

We previously found several amino acid differences in the sequence of CT288 among the different *C. trachomatis* strains (Table 3.3 in Chapter 3). Using Y2H, we next evaluated if CT288 from *C. trachomatis* ocular (serovar C) or urogenital (serovar E) strains could also interact with CCDC146. The plasmid used in the Y2H screen was constructed using $ct288_{\Delta N \Delta TMD}$ amplified from chromosomal DNA of *C. trachomatis* L2/434. We constructed similar plasmids, encoding Gal4BD-myc fusions of $CT288_{\Delta N \Delta TMD}$, where the *ct288* truncated gene was amplified from chromosomal DNA of *C. trachomatis* C/TW3 and E/Bour (Table A.2 and Table A.3 in Annexes). Immunoblotting analyses confirmed that all fusion proteins were expressed and migrated in a similar pattern in SDS-PAGE (Figure 4.4A). Y2HGold strains harboring combinations of plasmids encoding for the different Gal4BD-myc-$CT288_{\Delta N \Delta TMD}$ fusions, and for the Gal4AD-HA fusion of CCDC146, or empty plasmids, were plated in normal medium and in higher stringency medium. We could verify that $CT288_{\Delta N \Delta TMD}$ from each of the three *C. trachomatis* strains was able to interact with CCDC146 by Y2H (Figure 4.4B).

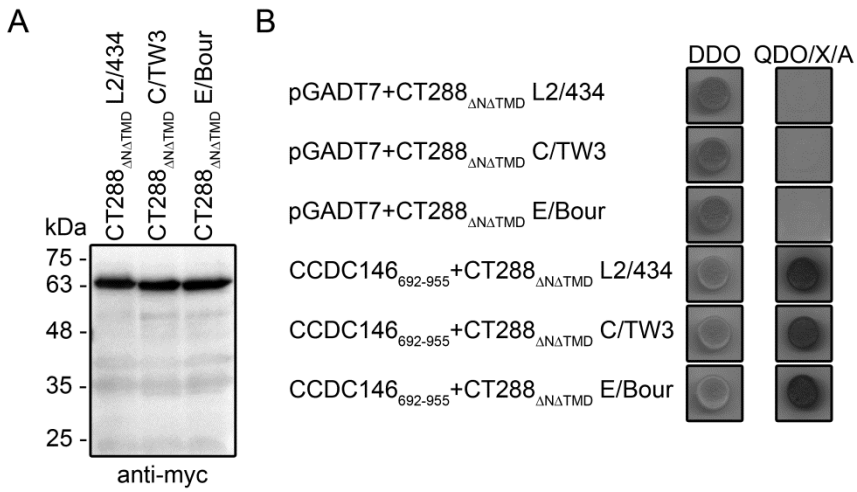


Figure 4.4 – CT288 from different *C. trachomatis* strains (L2/434, C/TW3, and E/Bour) interact with CCDC146 by Y2H. The sequence of *ct288ΔNΔTMD* from *C. trachomatis* strain L2/434 (the one used in the Y2H screen), C/TW3, and E/Bour were used to construct Gal4BD-myc-CT288$_{\Delta N\Delta TMD}$ fusions. (A) Immunoblot of yeast extracts after expression of the different Gal4BD-myc-CT288$_{\Delta N\Delta TMD}$ fusions from 3x10^7 yeast cells. (B) Interactions were verified by plating the Y2HGold strain harboring combinations of plasmids encoding for the different Gal4BD-myc-CT288$_{\Delta N\Delta TMD}$ fusions, and for the Gal4AD-HA fusion of CCDC146, or empty plasmids, in normal DDO medium and in higher stringency QDO/X/A medium.

4.4.4 Reciprocal co-immunoprecipitation of CT288 and CCDC146 after ectopic expression of the proteins in mammalian cells

To further test the interaction between CT288 and CCDC146, we performed co-IP assays after transient expression of the proteins in mammalian HEK293T cells. We constructed transfection plasmids encoding CT288, full-length (FL) CCDC146, or the C-terminal region of CCDC146 (amino acid residues 692-955) fused to the C-terminus of EGFP (EGFP-CT288$_{\Delta N\Delta TMD}$, EGFP-CCDC146$_{FL}$, and EGFP-CCDC146$_{692-955}$), or with a HA tag at their C-termini (CT288$_{\Delta N\Delta TMD}$-HA, CCDC146$_{FL}$-HA, and CCDC146$_{692-955}$-HA). These plasmids were then used to transiently transfect HEK293T cells for 24 h, using combinations of two plasmids, one encoding for CCDC146 and the other for CT288

(each with a different tag, EGFP or HA). As negative control, HEK293T cells were transiently transfected with plasmids encoding EGFP alone and HA-tagged CT288 or CCDC146. The cells were then lysed and EGFP or EGFP fusion proteins were immunoprecipitated. Subsequent immunoblotting analyses using anti-GFP and anti-HA antibodies revealed that $CT288_{\Delta N \Delta TMD}$-HA was pulled-down by the immunoprecipitation of EGFP-$CCDC146_{692\text{-}955}$ (Figure 4.5; middle panel) or EGFP-$CCDC146_{FL}$ (Figure 4.5; right panel), but not by the immunoprecipitation of EGFP alone (Figure 4.5; left panel). Similarly, $CCDC146_{692\text{-}955}$-HA (Figure 4.6A) or $CCDC146_{FL}$-HA (Figure 4.6B) were pulled-down by the immunoprecipitation of EGFP-$CT288_{\Delta N \Delta TMD}$ but not by EGFP alone (Figure 4.6).

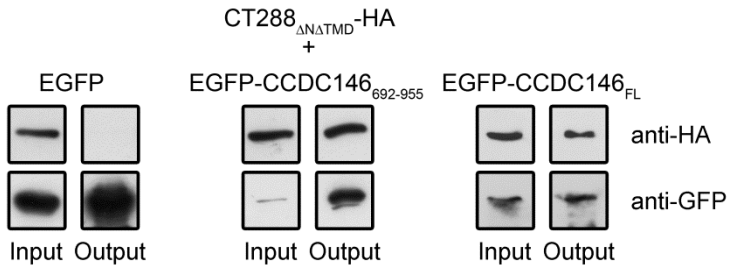


Figure 4.5 – Immunoprecipitation of transiently expressed EGFP-$CCDC146_{692\text{-}955}$ or EGFP-$CCDC146_{FL}$ is able to pull-down transiently expressed $CT288_{\Delta N \Delta TMD}$-HA in non-infected mammalian cells. Immunoblots of the supernatant fraction of HEK293T cells co-transfected for 24 h with combinations of plasmids encoding for EGFP, EGFP-$CCDC146_{FL}$, or EGFP-$CCDC146_{692\text{-}955}$, and a plasmid encoding for $CT288_{\Delta N \Delta TMD}$-HA (input – 0.5% loaded in SDS-PAGE) and of proteins immunoprecipitated by GFP antibodies using the GFP Trap kit (output – 20% loaded in SDS-PAGE).

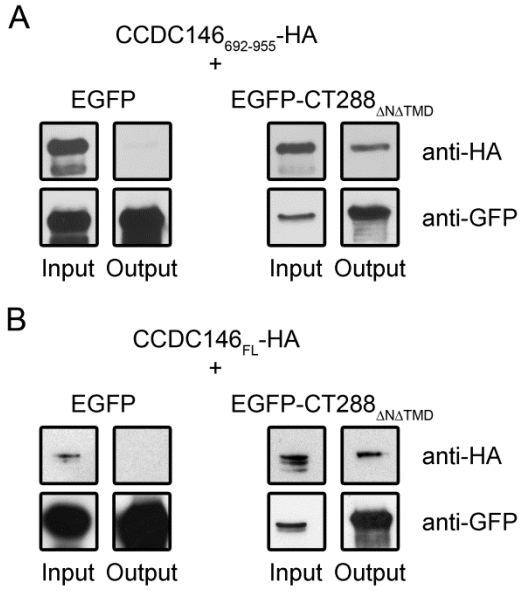


Figure 4.6 – Immunoprecipitation of transiently expressed EGFP-CT288$_{\Delta N \Delta TMD}$ is able to pull-down transiently expressed CCDC146$_{FL}$-HA or CCDC146$_{692-955}$-HA in non-infected mammalian cells. Immunoblots of the supernatant fraction of HEK293T cells co-transfected for 24 h with combinations of plasmids encoding for EGFP or EGFP-CT288$_{\Delta N \Delta TMD}$, and plasmids encoding for CCDC146$_{FL}$-HA (A) or CCDC146$_{692-955}$-HA (B) (input – 0.5% loaded in SDS-PAGE) and of proteins immunoprecipitated by GFP antibodies using the GFP Trap kit (output – 20% loaded in SDS-PAGE).

4.4.5 CT288 expressed by C. trachomatis during infection interacts with the C-terminal region of CCDC146

Next, we sought to test if CT288 and CCDC146 interact during infection. For this, we constructed a *C. trachomatis* L2/434 strain expressing full-length CT288 with a double HA tag at its C-terminus (CT288$_{FL}$-2HA). Immunoblotting analysis of whole cell extracts of HeLa 229 cells infected with the transformed and the parental *C. trachomatis* strain revealed that CT288$_{FL}$-2HA was expressed and migrated on SDS-PAGE as expected from the predicted molecular mass of 63 kDa (Figure 4.7A).

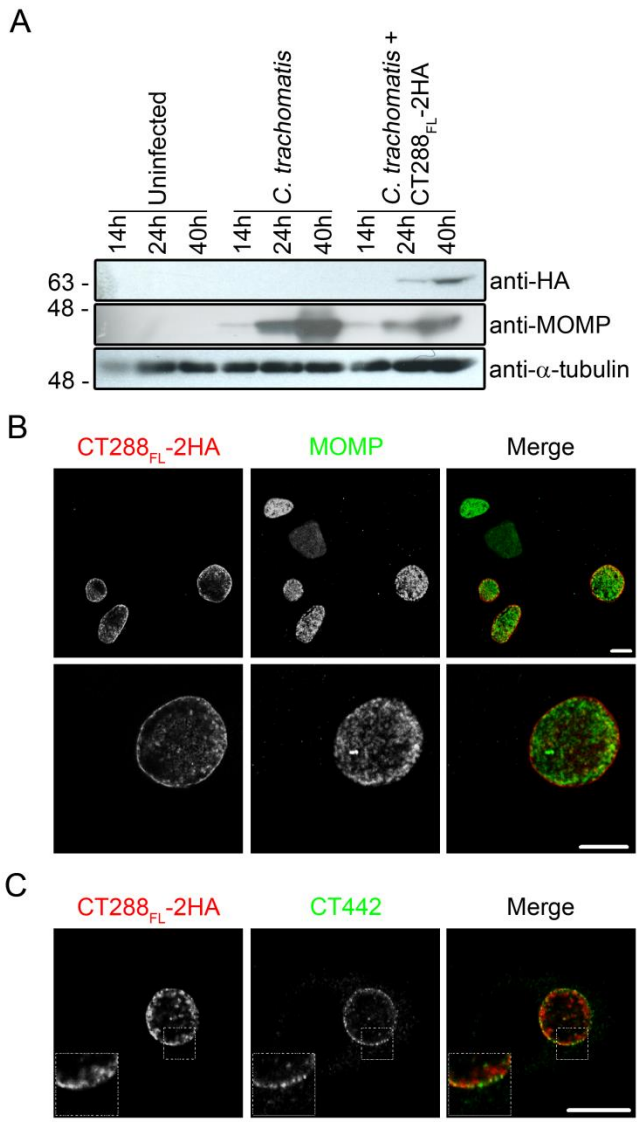


Figure 4.7 – CT288 localizes at the inclusion membrane in cells infected by *C. trachomatis* expressing CT288$_{FL}$-2HA. (A) Immunoblots of total extracts of HeLa 229 cells left uninfected or infected for 24 h with *C. trachomatis* L2/434 or *C. trachomatis* expressing CT288$_{FL}$-2HA. Proteins were detected with anti-HA, anti-MOMP, or anti-α-tubulin (as loading control) antibodies. (B) and (C) Hela 229 cells infected by *C. trachomatis* expressing CT288$_{FL}$-2HA for 24h were fixed with paraformaldehyde 4% (w/v), immunolabeled with anti-HA and anti-MOMP (B) or anti-HA and anti-CT442 (C) antibodies, and analyzed by immunofluorescence microscopy. In (B) the upper panel shows cells infected by transformed and untransformed *C. trachomatis* and the bottom panel shows one cell infected by transformed *C. trachomatis*. MOMP is the major outer membrane protein and localizes at the bacterial outer membrane, and CT442 is a known Inc protein. Scale bar is 10 µm.

Immunofluorescence microscopy analyses of HeLa 229 cells infected by *C. trachomatis* expressing $CT288_{FL}$-2HA with antibodies against HA and the major outer membrane protein (MOMP) of *C. trachomatis*, indicated that, as expected, CT288-2HA accumulates around the inclusion suggesting a localization at the bacterial vacuolar membrane (Figure 4.7B). Additional immunofluorescence microscopy analyses using antibodies against HA and Inc CT442 (Li *et al.*, 2008), confirmed the localization of $CT288_{FL}$-HA at the inclusion membrane (Figure 4.7C).

We then used *C. trachomatis* expressing $CT288_{FL}$-HA to infect for 24 h HEK293T cells that were also transiently transfected with plasmids encoding for EGFP or EGFP-$CCDC146_{692-955}$. Before lysis, 1% (w/v) paraformaldehyde was used as a cross-linking agent to maximize the probability of capturing the interaction between CT288 and CCDC146. The cells were then lysed and EGFP or EGFP-$CCDC146_{692-955}$ fusion proteins were immunoprecipitated. Subsequent immunoblotting analyses using anti-GFP and anti-HA antibodies revealed that $CT288_{FL}$-HA was pulled-down by the immunoprecipitation of EGFP-$CCDC146_{692-955}$ but not by EGFP alone (Figure 4.8).

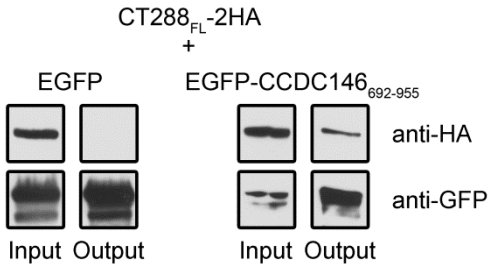


Figure 4.8 – Immunoprecipitation of ectopically expressed EGFP-$CCDC146_{FL}$ pulls-down $CT288_{FL}$-2HA in mammalian cells infected by *C. trachomatis* expressing $CT288_{FL}$-2HA. Immunoblots of the supernatant fraction of HEK293T cells infected for 24 h by *C. trachomatis* expressing $CT288_{FL}$-2HA and transfected with plasmids encoding for EGFP or EGFP-$CCDC146_{692-955}$ (input – 0.5% loaded in SDS-PAGE) and of proteins immunoprecipitated by GFP antibodies using the GFP Trap kit (output – 20% loaded in SDS-PAGE).

However, we could not replicate the same results when using full-length EGFP-CCDC146$_{FL}$ to pull-down CT288$_{FL}$-2HA from cells infected by *C. trachomatis* expressing CT288$_{FL}$-2HA (data not shown). Nevertheless, collectively these results support that CT288 interacts with CCDC146.

4.4.6 The subcellular localization of CCDC146 is altered in cells infected by C. trachomatis

A recent study indicated that endogenous CCDC146 concentrates at the centrosome (Firat-Karalar *et al.*, 2014). To analyze the localization of ectopically expressed CCDC146, HeLa 229 cells were transiently transfected for 24 h with plasmids encoding EGFP-CCDC146$_{FL}$ or CCDC146$_{FL}$-HA. The transfected cells were fixed with methanol and immunolabeled with anti-GFP and anti-γ-tubulin, or anti-HA and anti-γ-tubulin, and analyzed by immunofluorescence microscopy. Although we observed EGFP-CCDC146$_{FL}$ or CCDC146$_{FL}$-HA dispersed in the cytosol of HeLa 229 cells, we could also detect the proteins concentrated at the centrosome (Figure 4.9A and B; upper panels).

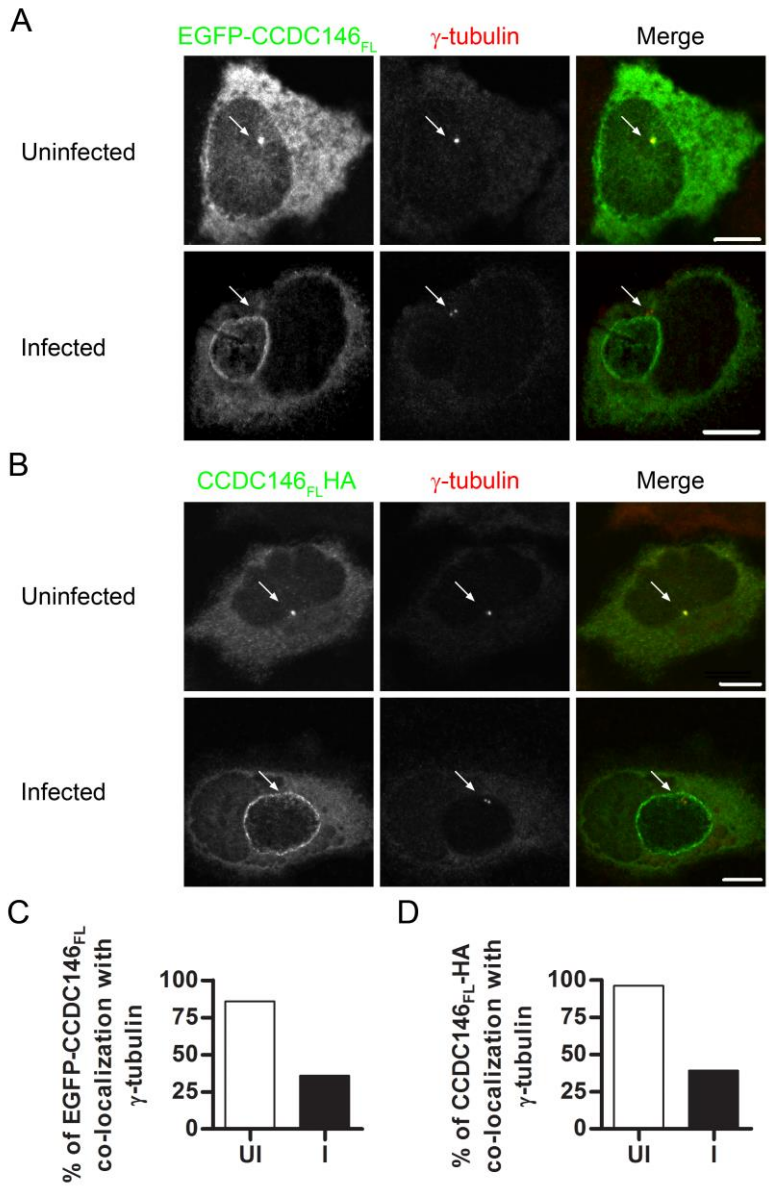


Figure 4.9 – CCDC146 loses its localization at the centrosome in mammalian cells infected with *C. trachomatis*. HeLa 229 cells transfected with plasmids encoding for EGFP-CCDC146$_{FL}$ (A) or CCDC146$_{FL}$-HA (B), were either left uninfected (UI) or infected with *C. trachomatis* L2/434 for 24h. The cells were fixed with methanol, immunolabeled with anti-GFP (A) or anti-HA (B), and anti-γ-tubulin antibodies, and analyzed by immunofluorescence microscopy. (C) and (D) Percentage (%) of co-localization of EGFP-CCDC146$_{FL}$ (C) or CCDC146$_{FL}$-HA (D), and anti-γ-tubulin. Data represents one single experiment where 100 cells were counted. The arrows in each panel highlight the centrosome. Scale bar is 10 μm.

Next, we investigated if the localization of ectopically expressed CCDC146 was altered due to infection by *C. trachomatis*. Hela 229 cells left uninfected or infected by *C. trachomatis* L2/434 for 24 h, and transiently expressing EGFP-CCDC146$_{FL}$ or CCDC146$_{FL}$-HA, were fixed with methanol, immunolabeled with antibodies against GFP or HA, and γ-tubulin, and analyzed by immunofluorescence microscopy. In infected cells, both EGFP-CCDC146$_{FL}$ and CCDC146$_{FL}$-HA apparently concentrated around the inclusion, suggesting a possible localization at the inclusion membrane (Figure 4.10A and B; lower panels). Furthermore, while in uninfected cells we could detect a co-localization of EGFP-CCDC146$_{FL}$ or CCDC146$_{FL}$-HA with γ-tubulin in about 90% of the cases, in infected cells we only detected co-localization of EGFP-CCDC146$_{FL}$ or CCDC146$_{FL}$-HA with γ-tubulin in about 40% of the cases (Figure 4.10C and D). This indicates that the subcellular localization of CCDC146 is altered during infection with *C. trachomatis*.

4.4.7 CCDC146 co-localizes with Inc proteins at the inclusion membrane in cells infected by C. trachomatis

We then addressed whether the ring-like accumulation of EGFP-CCDC146$_{FL}$ or CCDC146$_{FL}$-HA seemingly around the inclusion membrane corresponded to localization at the inclusion membrane. For this, HeLa 229 cells infected by *C. trachomatis* L2/434, and transiently expressing EGFP-CCDC146$_{FL}$ or CCDC146$_{FL}$-HA, were fixed with paraformaldehyde 4% (w/v) and immunolabeled with anti-GFP or anti-HA, and anti-MOMP (Figure 4.10A and B) or anti-CT442 (Figure 4.10C and D) antibodies. Immunofluorescence confocal microscopy analyses showed that EGFP-CCDC146$_{FL}$ and CCDC146$_{FL}$-HA were in fact surrounding the bacterial inclusion (Figure 4.10A and B). Furthermore, EGFP-CCC146$_{FL}$ and CCDC146$_{FL}$-HA co-localized with Inc CT442 around the inclusion (Figure 4.10C and D).

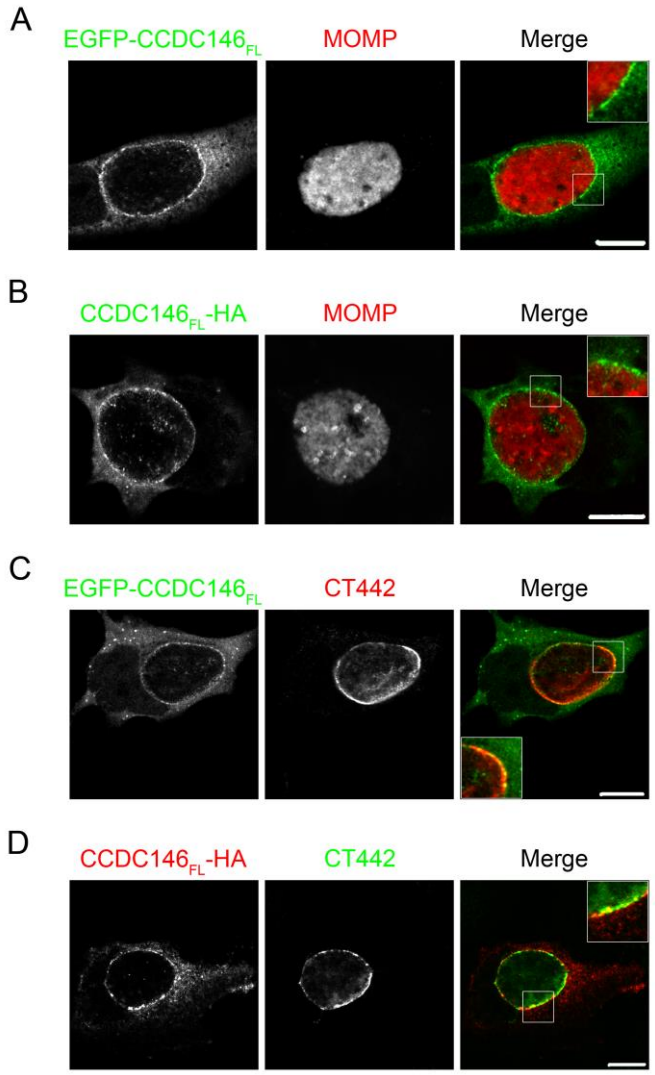


Figure 4.10 – CCDC146 localizes at the inclusion membrane in cells infected by *C. trachomatis*. HeLa 229 cells transfected with a plasmid encoding for EGFP-CCDC146$_{FL}$ (A and C), or CCDC146$_{FL}$-HA (B and D), were infected with *C. trachomatis* L2/434 for 24h. Cells were then fixed with paraformaldehyde 4% (w/v), immunolabeled with anti-GFP and anti-MOMP (A), or anti-CT442 (C), and with anti-HA and anti-MOMP (B), or anti-CT442 (D), and analyzed by immunofluorescence microscopy. Scale bar is 10 µm.

Finally, we examined the localization of ectopically expressed EGFP-CCDC146$_{FL}$ in HeLa 229 cells infected by *C. trachomatis* expressing CT288$_{FL}$-2HA (Figure 4.11). The infected and transfected cells were fixed with paraformaldehyde 4% (w/v) and immunolabeled with anti-HA and anti-GFP antibodies. In agreement with the previous analyses (Figure 4.9 and Figure 4.10), immunofluorescence microscopy analyses revealed that while EGFP is spread throughout the cytosol, EGFP-CCDC146$_{FL}$ co-localizes with CT288$_{FL}$-2HA at the inclusion membrane (Figure 4.11A). Altogether, these results indicate that CCDC146 and CT288 co-localize at the inclusion membrane during infection by *C. trachomatis*.

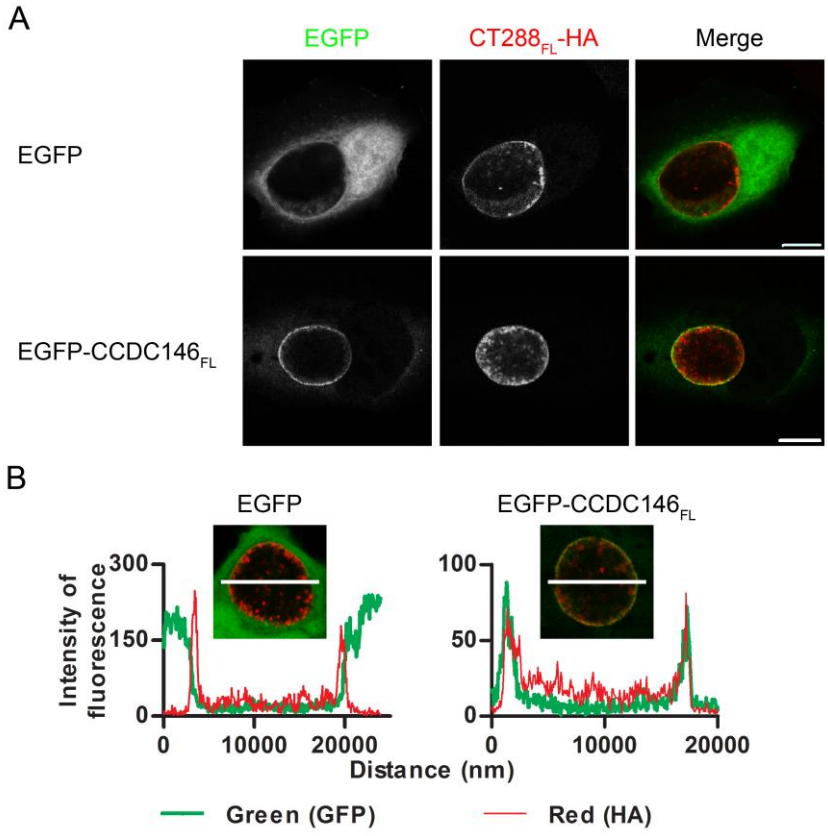


Figure 4.11 – CCDC146 co-localizes with CT288 at the inclusion membrane of HeLa 299 cells infected by *C. trachomatis* expressing CT288$_{FL}$-2HA. (A) HeLa 229 cells transfected with a plasmid encoding for EGFP or EGFP-CCDC146$_{FL}$, and infected by *C. trachomatis* expressing CT288$_{FL}$-2HA for 24h, were fixed with paraformaldehyde 4% (w/v), immunolabeled with anti-HA and anti-GFP antibodies, and analyzed by immunofluorescence microscopy. (B) Co-localization profile graphs of the immunofluorescence signal of GFP and anti-HA taken from confocal images. While in the case of EGFP, the signal does not overlap the signal of the CT288$_{FL}$-2HA, in the case of EGFP-CCDC146$_{FL}$, the signal completely overlaps the signal of the CT288$_{FL}$-2HA in the inclusion membrane. Scale bar is 10 µm.

4.5 Discussion

In this Chapter, we performed Y2H screens of a mammalian cDNA library to identify host cell candidate interacting partners of Inc proteins CT228, CT249, and CT288. We found an interaction between CT288 and the human centrosomal protein CCDC146. We further validated this interaction by co-IP experiments and described that during infection, CCDC146 co-localizes with CT288 at the inclusion membrane. Therefore, we unveil CCDC146 as a new host cell target for CT288, possibly involved in the association of *C. trachomatis* with the host cell centrosome, which occurs during infection with this bacterium.

Among the results for the other chlamydial Inc proteins (CT228 and CT249), at the same time we performed the Y2H screen with CT228, it was found by Lutter *et al.* that CT228 interacts with MYPT1 (Lutter *et al.*, 2013). In another study, among new candidate binding partners of Inc proteins identified by affinity purification-mass spectroscopy (AP-MS), the authors also validated the interaction between CT228 and MYPT1 (Mirrashidi *et al.*, 2015). As we could not replicate the results obtained by those authors and only found the interaction between CT228 and LCA5L, we decided to discard this result (Table 4.1).

For CT249, the most relevant Y2H results indicate that this protein might interact with USP11 (ubiquitin specific peptidase 11) and PEX1 (peroxisome biogenesis factor 1) (Table 4.1). USP11 has been proposed to regulate TNFα mediated NF-κB activation, and to participate in DNA damage repair functions within the BRCA2 pathway (Schoenfeld *et al.*, 2004; Sun *et al.*, 2010). PEX1 plays a role in the import of proteins into peroxisomes and peroxisome biogenesis (Tamura *et al.*, 1998). *C. trachomatis* is known to impair the DNA damage response, and is also known to hijack peroxisomes and utilize their enzymatic capacity to produce bacteria-specific phospholipids (Chumduri *et al.*, 2013;

Boncompain *et al.*, 2014). In the AP-MS protein-protein interaction network study, it was proposed that CT249 could interact with host proteins involved in cell division and lipid binding (Mirrashidi *et al.*, 2015). The overlap of functions found in these studies could indicate that the interactions found in our Y2H screen with CT249 might be relevant for the *C. trachomatis* infectious cycle. However, we did not further characterize these interactions.

The Y2H screen using CT288 revealed, among others, the interaction with CCDC146, which appeared 9 times in the Y2H screen, and with three other proteins related to chromosome partitioning and cell division (CENP-50, SMC3, and a subunit of PP2; Table 4.1). CCDC146 is a human protein of unknown function described to be localized at the centrosome, and possesses a region with similarity with the structural maintenance of chromosomes (SMC) domain (Strunnikov and Jessberger, 1999; Firat-Karalar *et al.*, 2014). This protein contains several coiled-coil regions, which are common among structural and motor proteins, and are usually involved in protein-protein interactions. The centromere protein 50 (CENP-50) belongs to the CENP-O class of proteins, suggested to be involved in the prevention of premature sister chromatin separation during the recovery from spindle damage (Minoshima *et al.*, 2005; Hori *et al.*, 2008). SMC3 possesses a SMC domain, and proteins possessing this domain are predicted to act as building blocks of chromosome structure, and to play a role in shaping chromatin and chromosomes (Strunnikov and Jessberger, 1999). PP2 (protein phosphatase 2) is a serine/threonine phosphatase that localizes at the centromeres and likely promotes chromosome–spindle interactions during cell division (Foley *et al.*, 2011). The possibility that these proteins (CENP-50, SMC3, and PP2) bind CT288, or interfere with its binding to CCDC146 remains to be analyzed.

We previously hypothesized that CT288 could be involved in the tissue tropism and type of infection associated with *C. trachomatis*. We found several amino acid differences in the sequence of CT288 among the different *C. trachomatis* strains that could separate the LGV strains (serovars L1-L3) from the ocular (serovars A-C) and urogenital (serovars D-K) strains (Table 3.3 in Chapter 3). Using Y2H, we investigated whether different versions of the CT288 protein, encoded by the DNA sequence of *C. trachomatis* C/TW3 (ocular strain), E/Bour (urogenital), and L2/434 (LGV strain; used in the Y2H screen), could interact with CCDC146. In the assay we performed all versions of CT288 interacted in a similar way with CCDC146 (Figure 4.4). However, the Y2H agar plating assay is not sensitive enough to detect possible small differences in the affinity of the interaction. In order to more accurately assess the binding of the different CT288 proteins to CCDC146, liquid α- or β-Galactosidase assays should be performed.

We determined by Y2H that the C-terminal region of CT288 (amino acids 292-563) is necessary for the interaction with the C-terminal region of CCDC146 (amino acids 692-955) (Figure 4.3). Moreover, this region of CT288 is believed to be on the host cytosol side of the inclusion membrane (Dehoux *et al.*, 2011). We also performed co-IP experiments to validate the interaction between CT288 and CCDC146. We observed that both the C-terminal and full-length CCDC146 were able to pull-down the truncated CT288 protein (without the N-terminal region and the two TMDs; $CT288_{\Delta N \Delta TMD}$) (Figure 4.5). In addition, $CT288_{\Delta N \Delta TMD}$ was able to pull-down both the CCDC146 C-terminal and full-length protein (Figure 4.6). In a more physiological approach, the C-terminal region of CCDC146 was able to precipitate full-length CT288 protein, expressed from the transformed *C. trachomatis* strain, in mammalian cells infected with this strain (Figure 4.8).

We used immunofluorescence microscopy to determine the fate of CCDC146 in mammalian cells infected with *C. trachomatis*. In uninfected cells, CCDC146 concentrates at the centrosome (Figure 4.9) (Firat-Karalar *et al.*, 2014). In *C. trachomatis* infected cells, we observed a shift in the localization of CCDC146 from the centrosomes to the inclusion membrane (Figure 4.9 and Figure 4.10). Moreover, we found that CCDC146 co-localizes with CT288 in the inclusion membrane in cells infected by *C. trachomatis* expressing CT288-2HA (Figure 4.11). Collectively, these results indicate that during infection by *C. trachomatis*, CCDC146 binds to CT288 and that this binding might lead to the recruitment of CCDC146 to the inclusion membrane.

CT288 is a 563 amino acids protein that (apart from the two bi-lobed hydrophobic domains) does not show significant identity to any known protein domains. CT288 is present in all *Chlamydia* but not in other *Chlamydia*-like organisms and in other bacteria. In contrast with other Inc proteins, in the AP-MS protein-protein interaction network study, no relevant results were found for this Inc protein (Mirrashidi *et al.*, 2015). CCDC146 is conserved among a large number of eukaryotes, including mammals, birds, and amphibians. This protein has 955 amino acids and the C-terminal region was recovered in the Y2H screen (amino acids 692 to 955).

The primary localization of CCDC146 is at the centrosome, that functions as the main microtubule organizing center (MTOC) in animal cells, and the coordinator of the cellular architecture and the bipolar spindles for DNA segregation during mitosis (Bettencourt-Dias and Glover, 2007). After invasion *Chlamydia* traffics along microtubules to the centrosome/MTOC in a dynein-dependent manner and maintains a tight association with it throughout the developmental cycle (Figure 4.12) (Grieshaber *et al.*, 2003; Grieshaber *et al.*, 2006; Richards *et al.*, 2013). This association disrupts normal centrosomal positioning and leads to a

loss of control of the centrosome duplication pathway that results in amplified centrosome numbers (Grieshaber *et al.*, 2006; Johnson *et al.*, 2009; Knowlton *et al.*, 2011). Infection by *Chlamydia* also affects host cell cytokinesis, by disrupting the cell cycle control, resulting in increased accumulation of host cell nuclei within cells (Greene and Zhong, 2003; Alzhanov *et al.*, 2009; Brown *et al.*, 2014). These events lead to the assembly of multipolar spindles that unequally distribute chromosomes during cell division (Grieshaber *et al.*, 2006; Alzhanov *et al.*, 2009; Brown *et al.*, 2014). As a consequence of chromosome instability, transformation and tumor development is likely to occur, and might explain the link between chlamydial infections and both human papilloma virus and cervical cancer (Koskela *et al.*, 2000; Sluder and Nordberg, 2004; Dahlström *et al.*, 2011; Weitzman and Weitzman, 2014; Nam *et al.*, 2015).

An earlier study reported that an active chlamydial transcription and translation machinery is necessary for *Chlamydia* to initiate trafficking to the MTOC (Scidmore *et al.*, 1996). This led to the suggestion that chlamydial proteins actively modify the inclusion membrane to facilitate this process. Due to their characteristic localization, Inc proteins are good candidates to be the factors that mediate these interactions. However, only a few Incs were associated with the centrosome. Earlier, it was reported that CT223 blocks cytokinesis *in vitro*, leading to multinucleation of the host cell (Alzhanov *et al.*, 2009). Recently, the same protein, now named inclusion protein acting on microtubules (IPAM), was described to interact with the centrosomal protein CEP170 (Dumoux *et al.*, 2015). IPAM and CEP170 co-localize at the inclusion membrane and interact to assemble microtubule superstructures necessary to maintain the inclusion shape and the success of the chlamydial infectious cycle (Figure 4.12) (Dumoux *et al.*, 2015). The Inc CT813/InaC was reported to bind to host ARF and 14-3-3 proteins, and

to promote the formation of F-actin structures around the inclusion, also important for the maintenance of the inclusion shape (Figure 4.12) (Kokes *et al.*, 2015). During infection, inclusion fusion is also an important event that occurs at the MTOC, and the host motor protein dynein and intact microtubules are key players in this process (Richards *et al.*, 2013).

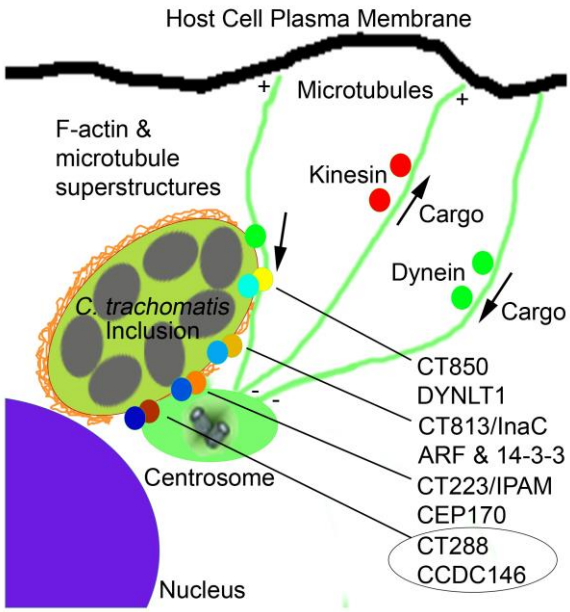


Figure 4.12 – Schematic model of the interactions between chlamydial proteins and centrosomal proteins. The centrosome coordinates cell architecture and regulate anterograde (kinesin-dependent) and retrograde (dynein-dependent) cargo transport. During infection, *Chlamydia* reside within a membranaceous compartment (inclusion), and traffic along the microtubules to the centrosome/MTOC in a dynein dependent manner, maintaining a tight association with the centrosome throughout the infectious cycle. Inc proteins were described to participate in these events. CT850 interacts with DYNLT1 to promote the positioning of the inclusion at the MTOC. CT223/IPAM interacts with the centrosomal protein CEP170, and CT813/InaC interacts with ARF and 14-3-3 proteins, promoting the assembly of microtubule superstructures necessary for inclusion shape and successful chlamydial infection. The interaction found in this Chapter between CT288 and the host cell centrosomal protein CCDC146 might also be functional and relevant for the inclusion dynamics with the centrosome.

The only chlamydial factor known to be involved in homotypic inclusion fusion is IncA, in which the encoded SNARE-like motif is believed to interact with host VAMP3 and VAMP8, and to promote fusion of compartments (Hackstadt *et al.*, 1999; Delevoye *et al.*, 2008). There is also a group of Incs (CT232/IncB, CT101, CT222, and CT850) that appear to localize in microdomains at the inclusion membrane, in regions that are enriched with active Src family kinases, which are associated with tyrosine phosphorylation processes (Mital *et al.*, 2010). These microdomains are localized at the point of contact of the centrosome with the inclusion membrane and might act as a scaffold for interactions controlling trafficking, positioning and integrity of the chlamydial inclusion (Mital *et al.*, 2010). Recently, CT850, one of the Inc proteins found in those microdomains, was described to interact with the dynein light chain DYNLT1 (Mital *et al.*, 2015). This interaction is necessary for a proper localization of the inclusion within the centrosome region (Figure 4.12) (Mital *et al.*, 2015). In the recent study where new host cell candidate binding partners of Inc proteins were identified by AP-MS, several Incs were associated with host cell cycle/division or centrosome processes (CT005, CT192, CT324, CT383, CT449, CT556, CT728, CT788, CT036, CT135, CT249, and CT846). However, these interactions still need further confirmation.

Although in this Chapter we could not verify the importance of the interaction between CT288 and CCDC146 in the context of the *C. trachomatis* developmental cycle, we hypothesize that this interaction might be functional during infection and important for the inclusion dynamics with the centrosome (Figure 4.12). Thus, we propose CCDC146 as a new host cell target protein for the Inc protein CT288. Further studies are needed to corroborate this hypothesis.

4.6 Acknowledgements

We are grateful to Guangming Zhong for the gift of the anti-CT442 antibody. The work reported in this Chapter was supported by Fundação para a Ciência e a Tecnologia (FCT) through grant PEst-OE/EQB/LA0004/2011, and by the European Commission through a Marie Curie European Re-integration Grant (PERG03-GA-2008-230954). Filipe Almeida was a recipient of a PhD fellowship from FCT (SFRH/BD/73545/2010).

4.7 References

Agaisse, H., and Derré, I. (2013) A *C. trachomatis* Cloning Vector and the Generation of *C. trachomatis* Strains Expressing Fluorescent Proteins under the Control of a *C. trachomatis* Promoter. *PLoS One* **8**.

Agaisse, H., and Derré, I. (2014) Expression of the effector protein IncD in *Chlamydia trachomatis* mediates recruitment of the lipid transfer protein CERT and the endoplasmic reticulum-resident protein VAPB to the inclusion membrane. *Infect Immun* **82**: 2037–2047.

Alzhanov, D.T., Weeks, S.K., Burnett, J.R., and Rockey, D.D. (2009) Cytokinesis is blocked in mammalian cells transfected with *Chlamydia trachomatis* gene CT223. *BMC Microbiol* **9**: 2.

Bettencourt-Dias, M., and Glover, D.M. (2007) Centrosome biogenesis and function: centrosomics brings new understanding. *Nat Rev Mol cell Biol* **8**: 451–463.

Betts, H.J., Wolf, K., and Fields, K.A. (2009) Effector protein modulation of host cells: examples in the *Chlamydia* spp. arsenal. *Curr Opin Microbiol* **12**: 81–87.

Boncompain, G., Müller, C., Meas-Yedid, V., Schmitt-Kopplin, P., Lazarow, P.B., and Subtil, A. (2014) The intracellular bacteria *Chlamydia* hijack peroxisomes and utilize their enzymatic capacity to produce bacteria-specific phospholipids. *PLoS One* **9**.

Brown, H.M., Knowlton, A.E., Snavely, E., Nguyen, B.D., Richards, T.S., and Grieshaber, S.S. (2014) Multinucleation during *C. trachomatis* infections is caused by the contribution of two effector pathways. *PLoS One* **9**: 1–14.

Chumduri, C., Gurumurthy, R.K., Zadora, P.K., Mi, Y., and Meyer, T.F. (2013) *Chlamydia* infection promotes host DNA damage and proliferation but impairs the DNA damage response. *Cell Host Microbe* **13**: 746–758.

Dahlström, L.A., Andersson, K., Luostarinen, T., Thoresen, S., Ögmundsdottír, H., Tryggvadottír, L., *et al.* (2011) Prospective seroepidemiologic study of human papillomavirus and other risk factors in cervical cancer. *Cancer Epidemiol Biomarkers Prev* **20**: 2541–2550.

Dehoux, P., Flores, R., Dauga, C., Zhong, G., and Subtil, A. (2011) Multi-genome identification and characterization of *Chlamydiae*-specific type III secretion substrates: the Inc proteins. *BMC Genomics* **12**: 109.

Delevoye, C., Nilges, M., Dehoux, P., Paumet, F., Perrinet, S., Dautry-Varsat, A., and Subtil, A. (2008) SNARE protein mimicry by an intracellular bacterium. *PLoS Pathog* **4**.

Derré, I., Swiss, R., and Agaisse, H. (2011) The lipid transfer protein CERT interacts with the *Chlamydia* inclusion protein IncD and participates to ER-*Chlamydia* inclusion membrane contact sites. *PLoS Pathog* **7**.

Dumoux, M., Menny, A., Delacour, D., and Hayward, R.D. (2015) A *Chlamydia* effector recruits CEP170 to reprogram host microtubule organization. *J Cell Sci* **44**.

Firat-Karalar, E.N., Sante, J., Elliott, S., and Stearns, T. (2014) Proteomic analysis of mammalian sperm cells identifies new components of the centrosome. *J Cell Sci* **3**: 4128–4133.

Foley, E. a., Maldonado, M., and Kapoor, T.M. (2011) Formation of stable attachments between kinetochores and microtubules depends on the B56-PP2A phosphatase. *Nat Cell Biol* **13**: 1265–1271.

Greene, W., and Zhong, G. (2003) Inhibition of host cell cytokinesis by *Chlamydia*

trachomatis infection. *J Infect* **47**: 45–51.

Grieshaber, S.S., Grieshaber, N. a., and Hackstadt, T. (2003) *Chlamydia trachomatis* uses host cell dynein to traffic to the microtubule-organizing center in a p50 dynamitin-independent process. *J Cell Sci* **116**: 3793–3802.

Grieshaber, S.S., Grieshaber, N. a., Miller, N.J., and Hackstadt, T. (2006) *Chlamydia trachomatis* causes centrosomal defects resulting in chromosomal segregation abnormalities. *Traffic* **7**: 940–949.

Hackstadt, T., Scidmore, M.A., Shaw, E.I., and Fischer, E.R. (1999) The *Chlamydia trachomatis* IncA protein is required for homotypic vesicle fusion. *Cell Microbiol* **1**: 119–130.

Hori, T., Okada, M., Maenaka, K., and Fukagawa, T. (2008) CENP-O Class Proteins Form a Stable Complex and Are Required for Proper Kinetochore Function. *Mol Biol Cell* **19**: 843–854.

Johnson, K. a., Tan, M., and Sütterlin, C. (2009) Centrosome abnormalities during a *Chlamydia trachomatis* infection are caused by dysregulation of the normal duplication pathway. *Cell Microbiol* **11**: 1064–1073.

Knowlton, A.E., Brown, H.M., Richards, T.S., Andreolas, L. a., Patel, R.K., and Grieshaber, S.S. (2011) *Chlamydia trachomatis* infection causes mitotic spindle pole defects independently from its effects on centrosome amplification. *Traffic* **12**: 854–866.

Kokes, M., Dunn, J.D., Granek, J.A., Nguyen, B.D., Barker, J.R., Valdivia, R.H., and Bastidas, R.J. (2015) Integrating Chemical Mutagenesis and Whole-Genome Sequencing as a Platform for Forward and Reverse Genetic Analysis of *Chlamydia*. *Cell Host Microbe* 1–10.

Koskela, P., Anttila, T., Bjørge, T., Brunsvig, A., Dillner, J., Hakama, M., *et al.* (2000) *Chlamydia trachomatis* infection as a risk factor for invasive cervical cancer. *Int J Cancer* **85**: 35–39.

Li, Z., Chen, C., Chen, D., Wu, Y., Zhong, Y., and Zhong, G. (2008) Characterization of fifty putative inclusion membrane proteins encoded in the *Chlamydia trachomatis* genome. *Infect Immun* **76**: 2746–2757.

Lutter, E.I., Barger, A.C., Nair, V., and Hackstadt, T. (2013) *Chlamydia trachomatis* Inclusion Membrane Protein CT228 Recruits Elements of the Myosin Phosphatase Pathway to Regulate Release Mechanisms. *Cell Rep* **3**: 1921–1931.

Minoshima, Y., Hori, T., Okada, M., Kimura, H., Haraguchi, T., Hiraoka, Y., *et al.* (2005) The constitutive centromere component CENP-50 is required for recovery from spindle damage. *Mol Cell Biol* **25**: 10315–10328.

Mirrashidi, K.M., Elwell, C.A., Verschueren, E., Johnson, J.R., Frando, A., Dollen, J. Von, *et al.* (2015) Global Mapping of the Inc-Human Interactome Reveals that Retromer Restricts *Chlamydia* Infection. *Cell Host Microbe* **18**: 109–121.

Mital, J., Lutter, E.I., Barger, A.C., Dooley, C.A., and Hackstadt, T. (2015) *Chlamydia trachomatis* inclusion membrane protein CT850 interacts with the dynein light chain DYNLT1 (Tctex1). *Biochem Biophys Res Commun* **1**: 1–6.

Mital, J., Miller, N.J., Fischer, E.R., and Hackstadt, T. (2010) Specific chlamydial inclusion membrane proteins associate with active Src family kinases in microdomains that interact with the host microtubule network. *Cell Microbiol* **12**: 1235–1249.

Nam, H.-J., Naylor, R.M., and Deursen, J.M. van (2015) Centrosome dynamics as a source of chromosomal instability. *Trends Cell Biol* **25**: 65–73.

Richards, T.S., Knowlton, A.E., and Grieshaber, S.S. (2013) *Chlamydia trachomatis* homotypic inclusion fusion is promoted by host microtubule trafficking. *BMC Microbiol* **13**: 185.

Rzomp, K.A., Moorhead, A.R., and Scidmore, M.A. (2006) The GTPase Rab4 interacts with *Chlamydia trachomatis* inclusion membrane protein CT229. *Infect Immun* **74**: 5362–5373.

Schoenfeld, A.R., Apgar, S., Dolios, G., Wang, R., and Aaronson, S. a (2004) BRCA2 is ubiquitinated in vivo and interacts with USP11, a deubiquitinating enzyme that exhibits prosurvival function in the cellular response to DNA damage. *Mol Cell Biol* **24**: 7444–7455.

Scidmore, M.A. (2005) Culture and Cultivation of *Chlamydia trachomatis*. *Curr Protoc Microbiol* 1–25.

Scidmore, M.A., and Hackstadt, T. (2001) Mammalian 14-3-3-beta associates with the *Chlamydia trachomatis* inclusion membrane via its interaction with IncG. *Mol Microbiol* **39**: 1638–1650.

Scidmore, M.A., Rockey, D.D., Fischer, E.R., Heinzen, R.A., and Hackstadt, T. (1996) Vesicular interactions of the *Chlamydia trachomatis* inclusion are determined by chlamydial early protein synthesis rather than route of entry. *Infect Immun* **64**: 5366–5372.

Sluder, G., and Nordberg, J.J. (2004) The good, the bad and the ugly: The practical consequences of centrosome amplification. *Curr Opin Cell Biol* **16**: 49–54.

Strunnikov, A. V., and Jessberger, R. (1999) Structural maintenance of chromosomes (SMC) proteins: Conserved molecular properties for multiple biological functions. *Eur J Biochem* **263**: 6–13.

Sun, W., Tan, X., Shi, Y., Xu, G., Mao, R., Gu, X., *et al.* (2010) USP11 negatively regulates TNFalfa-induced NF-B activation by targeting on IκBα. *Cell Signal* **22**: 386–394.

Tamura, S., Okumoto, K., Toyama, R., Shimozawa, N., Tsukamoto, T., Suzuki, Y., *et al.* (1998) Human PEX1 cloned by functional complementation on a CHO cell mutant is responsible for peroxisome-deficient Zellweger syndrome of complementation group I. *Proc Natl Acad Sci U S A* **95**: 4350–4355.

Valdivia, R.H. (2008) *Chlamydia* effector proteins and new insights into chlamydial cellular microbiology. *Curr Opin Microbiol* **11**: 53–59.

Weitzman, M.D., and Weitzman, J.B. (2014) What's the damage? The impact of pathogens on pathways that maintain host genome integrity. *Cell Host Microbe* **15**: 283–294.

CHAPTER 5 - Conclusions and future perspectives

The author of this dissertation wrote the whole Chapter, based on the referred bibliography and the results described in Chapters 2 to 4.

The work developed in this thesis provided novel insights into the mechanisms involved in infections by *Chlamydia trachomatis*. In particular: i) we provide additional and thorough support for inclusion membrane (Inc) proteins being type III secretion (T3S) substrates (Chapter 2); ii) we found evidence for a role of a subset of Inc proteins in the tissue tropism and type of infections associated with *C. trachomatis* (Chapter 3); and iii) we discovered an interaction between the Inc CT288 and the host centrosomal protein coiled coil domain containing protein 146 (CCDC146), which during infection of host cells by *C. trachomatis* we found to localize at the inclusion membrane (Chapter 4).

5.1 T3S of Inc proteins

The first reports of Inc proteins were given in 1995 and 1998 (Rockey *et al.*, 1995; Bannantine *et al.*, 1998). By using sera from guinea-pigs that recovered from ocular infections with *C. psittaci*, three proteins were isolated and characterized that localized at the inclusion membrane of infected cells. Due to their localization, the genes encoding these proteins were named *incA*, *incB*, and *incC*, for inclusion membrane proteins A, B, and C, respectively. Amino acid sequence and secondary structure analyses of these proteins revealed little similarity between them, except for a long hydrophobic region indicative of a membrane spanning domain (Rockey *et al.*, 1995; Bannantine *et al.*, 1998).

With the first chlamydial genomes available, the use of bioinformatics allowed the search for other proteins encoded in the genome possessing the characteristic bi-lobed hydrophobic motif of at least 50 amino acids (Bannantine *et al.*, 2000). Several candidate Inc proteins (~70) were discovered by this approach; however, an important question arose from these findings, which was how do these proteins could reach the inclusion membrane? As it was already known that *C. trachomatis*

possesses a functional type III secretion system (T3SSs) (Fields and Hackstadt, 2000), it was hypothesized that Inc proteins could be translocated from the bacteria into the host cell by this mechanism. Subtil and colleagues found that the N-terminal of *C. trachomatis* and *C. pneumoniae* Inc proteins was sufficient for driving T3S of a fusion protein by the *S. flexneri* T3SS (Subtil *et al.*, 2001; Subtil *et al.*, 2005). Later, in a comprehensive bioinformatics study, Dehoux and colleagues identified and analyzed all candidate Inc proteins in the available genomes of seven chlamydial species (discovering a total of 537 putative Incs) (Dehoux *et al.*, 2011). However, the presence of T3S signals was not tested for all of them. Other authors tried to determine the localization of all Inc proteins by generating antibodies against a set of ~50 recombinant Inc proteins, and they could confirm the localization at the inclusion membrane for ~20 Incs (Li *et al.*, 2008). Subsequent studies increased the number of validated Inc proteins (possessing the bi-lobed hydrophobic motif and localizing at the inclusion membrane) to ~25 (Cortes *et al.*, 2007; Li *et al.*, 2008; Mital *et al.*, 2010; Dehoux *et al.*, 2011; Flores and Zhong, 2015).

As a follow up of these studies, we asked if the presence of a T3S signal could be further evidence (in addition to the bi-lobed hydrophobic domain) that an Inc protein localizes at the inclusion membrane. Focusing on a set of 48 *C. trachomatis* known and putative Inc proteins (Li *et al.*, 2008), we identified T3S signals in the N-terminal of about 80% of the Incs. However, we could not identify a T3S signal in a subset of known Inc proteins, demonstrating that definitive localization at the inclusion membrane needs confirmation by immunolocalization studies in *Chlamydia*-infected cells. These studies are now facilitated by the recent advances in genetic manipulation of *Chlamydia*. In fact, *C. trachomatis* strains expressing recombinant Inc proteins with a FLAG epitope tag were constructed (Bauler and Hackstadt, 2014; Weber *et al.*,

2015). Subsequent immunolocalization studies of these recombinant proteins in host cells infected by *C. trachomatis* encoding FLAG-tagged Inc proteins allowed the identification of 10 additional Inc proteins localizing at the inclusion membrane (CT006, CT134, CT135, CT179, CT192, CT224, CT227, CT345, CT383, and CT449) (Weber *et al.*, 2015). Overall, our results were in agreement with these observations, as we identified T3S signals in 7 of these 10 Inc proteins (CT135, CT92, CT225, CT227, CT345, CT383, and CT449). Furthermore, for 4 Inc proteins (CT483, CT484, CT565, and CT728), in which we did not identify a T3S signal, those authors also did not detect localization of the recombinant FLAG-tagged proteins at the inclusion membrane (Weber *et al.*, 2015). However, for 6 Inc proteins (CT036, CT058, CT195, CT214, CT365, and CT789), in which we identified T3S signals, those authors did not detect localization of the recombinant FLAG-tagged proteins at the inclusion membrane (Weber *et al.*, 2015).

While our experiments and those previously done in other laboratories (Fields *et al.*, 2003; Subtil *et al.*, 2005; Dehoux *et al.*, 2011) clearly support that the majority of Inc proteins are translocated into the host cell by a T3SS, at a mechanistic level it remains to be understood how they insert into the inclusion membrane after translocation. This presumably involves the characteristic bi-lobed hydrophobic domain but the molecular details are unknown.

5.2 A role of Inc proteins in tissue tropism and type of infection associated with *C. trachomatis*

The epitheliotropic (ocular and genital) and lymphogranuloma venereum (LGV) strains share > 98% of identity at the DNA level (Stephens *et al.*, 1998; Carlson *et al.*, 2005; Thomson *et al.*, 2008; Seth-Smith *et al.*, 2009; Jeffrey *et al.*, 2010; Unemo *et al.*, 2010; Somboonna *et al.*, 2011;

Harris *et al.*, 2012). Hence, the different tissue tropisms and disease outcomes associated to infections by *C. trachomatis* are likely a consequence of small-scale mutational changes or of variations in gene expression. A clear example is the loss of tryptophan synthase function in the strains associated with ocular infections. All ocular strains have a defective *trpB* gene within the tryptophan (*trpBA*) operon, while urogenital and LGV strains have an intact copy of the gene (Allan C. Shaw *et al.*, 2000; Fehlner-Gardiner *et al.*, 2002; Caldwell *et al.*, 2003). Other examples include variations in a few genes located in the plasticity zone (the *Chlamydia* cytotoxin gene and genes encoding phospholipase D-like proteins), TARP and Pmps (Carlson *et al.*, 2004; Gomes *et al.*, 2006; Nelson *et al.*, 2006; Lutter *et al.*, 2010).

The first report of Inc proteins possibly involved in the tissue tropism and different types of infection was given by Brunelle and colleagues in 2004 (Brunelle *et al.*, 2004). By using DNA microarray analyses, the authors compared the genome of serovar D to that of other 14 serovars, and found 4 *inc* genes (*ct116/incE*, *ct223*, *ct288*, and *ct618*) among 11 genes with differences associated with LGV strains. Two other groups performed phylogenetic and molecular evolutionary analyses in the genomes of strains representative of each disease group and also found that *inc* genes and Inc proteins clearly separate epitheliotropic from LGV strains (Nunes *et al.*, 2008; Thomson *et al.*, 2008). Three different groups analyzed the temporal gene expression of chlamydial genes during the developmental cycle (E. I. Shaw *et al.*, 2000; Belland *et al.*, 2003; Nicholson *et al.*, 2003). These studies were performed only with serovars D and L2, thus, did not cover all the three disease groups. Furthermore, different methodologies were used and therefore, definitive conclusions about the variations in gene expression identified in these studies could not be made. Given that the genes encoding Inc proteins comprise ~10% of the genome (Dehoux *et al.*, 2011), it was suggested

that *inc* genes or Inc proteins could contribute to *C. trachomatis* tissue tropism, and in particular to disseminating infections caused by LGV strains. However, a thorough and detailed analysis of all putative *inc* genes and Inc proteins in all 15 serovars had never been pursued.

Focusing on a set of 48 known and putative *C. trachomatis* Inc proteins (Li *et al.*, 2008), we could identify small differences in the amino acid sequence of a subgroup of Inc proteins and in the expression of *inc* genes that correlate with the specific tissue tropism and type of infection associated with LGV strains. Recently, a similar study was published, where the authors used the whole genome (~900 genes) of ~50 *C. trachomatis* strains for the phylogenetics and molecular evolutionary analyses (Ferreira *et al.*, 2015). In agreement with our work, the evolutionary analyses revealed that the genes with significant $d_N/d_S > 1$ were those mainly encoding Inc protein and other T3SS effectors, supporting the hypothesis that proteins directly involved in host cell-pathogen interactions during infection are subjected to selective pressures (Ferreira *et al.*, 2015).

Most studies on *C. trachomatis* tissue tropism are based on bioinformatics and the relevance of the correlations found remains to be directly analyzed. With the ongoing developments in the genetic manipulation of *C. trachomatis* and with the animal models available (the pig-tailed macaque for eye infections and the *C. muridarum* mouse model for genital and LGV infections), it is conceivably possible that in the future the exact role of specific *C. trachomatis* genes, such as *inc* genes, on tissue tropism can be demonstrated.

The thorough analysis of *inc* gene expression we conducted helped to classify *inc* genes according to their levels and profiles of expression. These data could be used to gain insights into the mechanisms of regulation of the expression of *inc* genes. Assuming that expression of

inc genes is controlled by σ^{66} (preliminary results suggest strong -10 and -35 elements in several *inc* genes), then other mechanisms to differentially regulate the temporal classes of genes must be present. Additionally, multiple mechanisms of transcriptional regulation might occur, such as DNA supercoiling or late regulation by the repressor early upstream ORF (EUO) and σ^{28} (Cheng and Tan, 2012; Rosario and Tan, 2012). Bioinformatics analyses of the putative promoter regions could be performed in an attempt to find possible binding sites for transcription regulators. This knowledge could then be used to identify transcriptional regulators either by using *Escherichia coli* as a heterologous host (Abdelrahman *et al.*, 2011), or by constructing libraries of recombinant proteins of putative transcriptional regulators in *C. trachomatis* using transformation methods (Weber *et al.*, 2015).

5.3 Binding between Inc protein CT288 and the host cell centrosomal protein CCDC146

We used yeast-two hybrid screening to find host cell interacting partners of three Inc proteins (CT228, CT249, CT288) hypothesized to contribute to the tissue tropism and type of infection associated with *C. trachomatis* (as shown in Chapter 3), with the expectation that the host cell interacting proteins found could suggest a function for these Inc proteins. We focused on the interaction between CT288 and CCDC146, a human protein of unknown function that localizes at the centrosome, and more specifically at the centrioles (Firat-Karalar *et al.*, 2014). We validated the interaction between CT288 and CCDC146 by co-IP and described that CCDC146 co-localizes with CT288 in the inclusion membrane during infection by *C. trachomatis*. The centrosome is an important structure in eukaryotic cells that functions as coordinator of cellular architecture (Bettencourt-Dias and

Glover, 2007), and it is known that *C. trachomatis* maintains a tight association with it throughout the infectious cycle (Grieshaber *et al.*, 2003; Grieshaber *et al.*, 2006; Richards *et al.*, 2013). In addition, a few centrosomal proteins were already identified as host cell targets of Inc proteins, indicating that Inc proteins are involved in the cross-talk between the inclusion and the centrosome (Mital *et al.*, 2010; Dumoux *et al.*, 2015; Kokes *et al.*, 2015). Therefore, the interaction found in this Chapter between CT288 and the host cell centrosomal protein CCDC146 might be functional and relevant during *C. trachomatis* infections.

Due to time restrictions, it was not possible to evaluate the importance of the interaction between CT288 and CCDC146 during the developmental cycle of *C. trachomatis*. Future obvious experiments include the generation of a *C. trachomatis ct288* mutant strain, and to silence the expression of the *ccdc146* in host cells. We could then evaluate the importance of these two genes during infection by *C. trachomatis* of host cells.

In the Y2H screenings, while we discarded the results obtained with CT228, we found several candidates for CT249 and CT288 that were not further investigated. For CT249, the interactions with a protein involved in DNA damage repair functions and with a protein involved in peroxisome biogenesis are of particular interest, as *C. trachomatis* is known to interfere with these processes during infection (Chumduri *et al.*, 2013; Boncompain *et al.*, 2014). For CT288, additional proteins related with centrosome functions (CENP-50, SMC3, and PP2) were also found in the Y2H screen using CT288 as bait. It will be interesting to address if there is any link between these proteins and CCDC146.

5.4 References

Abdelrahman, Y.M., Rose, L. a., and Belland, R.J. (2011) Developmental expression of non-coding RNAs in *Chlamydia trachomatis* during normal and persistent growth. *Nucleic Acids Res* **39**: 1843–1854.

Bannantine, J.P., Griffiths, R.S., Viratyosin, W., Brown, W.J., and Rockey, D.D. (2000) A secondary structure motif predictive of protein localization to the chlamydial inclusion membrane. *Cell Microbiol* **2**: 35–47.

Bannantine, J.P., Rockey, D.D., and Hackstadt, T. (1998) Tandem genes of *Chlamydia psittaci* that encode proteins localized to the inclusion membrane. *Mol Microbiol* **28**: 1017–1026.

Bauler, L.D., and Hackstadt, T. (2014) Expression and Targeting of secreted proteins from *Chlamydia trachomatis*. *J Bacteriol* **196**: 1325–1334.

Belland, R.J., Zhong, G., Dean, D., Hogan, D., Sturdevant, D.E., Sharma, J., *et al.* (2003) Genomic transcriptional profiling of the developmental cycle of *Chlamydia trachomatis*. *Proc Natl Acad Sci U S A* **100**: 8478–8483.

Bettencourt-Dias, M., and Glover, D.M. (2007) Centrosome biogenesis and function: centrosomics brings new understanding. *Nat Rev Mol cell Biol* **8**: 451–463.

Boncompain, G., Müller, C., Meas-Yedid, V., Schmitt-Kopplin, P., Lazarow, P.B., and Subtil, A. (2014) The intracellular bacteria *Chlamydia* hijack peroxisomes and utilize their enzymatic capacity to produce bacteria-specific phospholipids. *PLoS One* **9**.

Brunelle, B.W., Nicholson, T.L., and Stephens, R.S. (2004) Microarray-based genomic surveying of gene polymorphisms in *Chlamydia trachomatis*. *Genome Biol* **5**: R42.

Caldwell, H.D., Wood, H., Dean, D., Bailey, R., Jones, R.B., Mabey, D., *et al.* (2003) Polymorphisms in *Chlamydia trachomatis* tryptophan synthase genes differentiate between genital and ocular isolates. *J Clin Invest* **111**: 1757–1769.

Carlson, J.H., Hughes, S., Hogan, D., Cieplak, G., Sturdevant, D.E., McClarty, G., *et al.* (2004) Polymorphisms in the *Chlamydia trachomatis* cytotoxin locus associated with ocular and genital isolates. *Infect Immun* **72**: 7063–7072.

Carlson, J.H., Porcella, S.F., Mcclarty, G., and Caldwell, H.D. (2005) Comparative Genomic Analysis of *Chlamydia trachomatis* Oculotropic and Genitotropic Strains. *Infect Immun* **73**: 6407–6418.

Cheng, E., and Tan, M. (2012) Differential effects of DNA supercoiling on *Chlamydia* early promoters correlate with expression patterns in midcycle. *J Bacteriol* **194**: 3109–3115.

Chumduri, C., Gurumurthy, R.K., Zadora, P.K., Mi, Y., and Meyer, T.F. (2013) *Chlamydia* infection promotes host DNA damage and proliferation but impairs the DNA damage response. *Cell Host Microbe* **13**: 746–758.

Cortes, C., Rzomp, K.A., Tvinnereim, A., Scidmore, M.A., and Wizel, B. (2007) *Chlamydia pneumoniae* inclusion membrane protein Cpn0585 interacts with multiple Rab GTPases. *Infect Immun* **75**: 5586–96.

Dehoux, P., Flores, R., Dauga, C., Zhong, G., and Subtil, A. (2011) Multi-genome identification and characterization of *Chlamydiae*-specific type III secretion substrates: the Inc proteins. *BMC Genomics* **12**: 109.

Dumoux, M., Menny, A., Delacour, D., and Hayward, R.D. (2015) A *Chlamydia* effector recruits CEP170 to reprogram host microtubule organization. *J Cell Sci* **44**.

Fehlner-Gardiner, C., Roshick, C., Carlson, J.H., Hughes, S., Belland, R.J., Caldwell, H.D., and McClarty, G. (2002) Molecular basis defining human *Chlamydia trachomatis* tissue tropism: A possible role for tryptophan synthase. *J Biol Chem* **277**: 26893–26903.

Ferreira, R., Antelo, M., Nunes, A., Borges, V., Damião, V., Borrego, M.J., and Gomes, J.P. (2015) *In Silico* Scrutiny of Genes Revealing Phylogenetic Congruence with Clinical Prevalence or Tropism Properties of *Chlamydia trachomatis* Strains. **5**: 9–19.

Fields, K.A., and Hackstadt, T. (2000) Evidence for the secretion of *Chlamydia trachomatis* CopN by a type III secretion mechanism. *Mol Microbiol* **38**: 1048–1060.

Fields, K.A., Mead, D.J., Dooley, C.A., and Hackstadt, T. (2003) *Chlamydia trachomatis* type III secretion: Evidence for a functional apparatus during early-cycle development. *Mol Microbiol* **48**: 671–683.

Firat-Karalar, E.N., Sante, J., Elliott, S., and Stearns, T. (2014) Proteomic analysis of mammalian sperm cells identifies new components of the centrosome. *J Cell Sci* **3**: 4128–4133.

Flores, R., and Zhong, G. (2015) The *Chlamydia pneumoniae* Inclusion Membrane Protein Cpn1027 Interacts with Host Cell Wnt Signaling Pathway Regulator Cytoplasmic Activation/Proliferation-Associated Protein 2 (Caprin2). *PLoS One* **10**: e0127909.

Gomes, J.P., Nunes, A., Bruno, W.J., Borrego, M.J., Florindo, C., and Dean, D. (2006) Polymorphisms in the nine polymorphic membrane proteins of *Chlamydia trachomatis* across all serovars: Evidence for serovar Da recombination and correlation with tissue tropism. *J Bacteriol* **188**: 275–286.

Grieshaber, S.S., Grieshaber, N. a., and Hackstadt, T. (2003) *Chlamydia trachomatis* uses host cell dynein to traffic to the microtubule-organizing center in a p50 dynamitin-independent process. *J Cell Sci* **116**: 3793–3802.

Grieshaber, S.S., Grieshaber, N. a., Miller, N.J., and Hackstadt, T. (2006) *Chlamydia trachomatis* causes centrosomal defects resulting in chromosomal segregation abnormalities. *Traffic* **7**: 940–949.

Harris, S.R., Clarke, I.N., Seth-Smith, H.M.B., Solomon, A.W., Cutcliffe, L.T., Marsh, P., *et al.* (2012) Whole-genome analysis of diverse *Chlamydia trachomatis* strains identifies phylogenetic relationships masked by current clinical typing. *Nat Genet* **44**: 413–419.

Jeffrey, B.M., Suchland, R.J., Quinn, K.L., Davidson, J.R., Stamm, W.E., and Rockey, D.D. (2010) Genome sequencing of recent clinical *Chlamydia trachomatis* strains identifies loci associated with tissue tropism and regions of apparent recombination. *Infect Immun* **78**: 2544–2553.

Kokes, M., Dunn, J.D., Granek, J.A., Nguyen, B.D., Barker, J.R., Valdivia, R.H., and Bastidas, R.J. (2015) Integrating Chemical Mutagenesis and Whole-Genome Sequencing as a Platform for Forward and Reverse Genetic Analysis of *Chlamydia*. *Cell Host Microbe* 1–10.

Li, Z., Chen, C., Chen, D., Wu, Y., Zhong, Y., and Zhong, G. (2008) Characterization of fifty putative inclusion membrane proteins encoded in the *Chlamydia trachomatis* genome. *Infect Immun* **76**: 2746–2757.

Lutter, E.I., Bonner, C., Holland, M.J., Suchland, R.J., Stamm, W.E., Jewett, T.J., *et al.* (2010) Phylogenetic analysis of *Chlamydia trachomatis* tarp and correlation with clinical phenotype. *Infect Immun* **78**: 3678–3688.

Mital, J., Miller, N.J., Fischer, E.R., and Hackstadt, T. (2010) Specific chlamydial inclusion membrane proteins associate with active Src family kinases in microdomains that interact with the host microtubule network. *Cell Microbiol* **12**: 1235–1249.

Nelson, D.E., Dean, D., Taylor, L.D., Dorward, D.W., Goheen, M.M., and Caldwell, H.D. (2006) Inhibition of *Chlamydiae* by primary alcohols correlates with the strain-specific complement of plasticity zone phospholipase D genes. *Infect Immun* **74**: 73–80.

Nicholson, T.L., Olinger, L., Chong, K., Schoolnik, G., and Stephens, R.S. (2003) Global stage-specific gene regulation during the developmental cycle of *Chlamydia trachomatis*. *J Bacteriol* **185**: 3179–3189.

Nunes, A., Nogueira, P.J., Borrego, M.J., and Gomes, J.P. (2008) *Chlamydia trachomatis* diversity viewed as a tissue-specific coevolutionary arms race. *Genome Biol* **9**: R153.

Richards, T.S., Knowlton, A.E., and Grieshaber, S.S. (2013) *Chlamydia trachomatis* homotypic inclusion fusion is promoted by host microtubule trafficking. *BMC Microbiol* **13**: 185.

Rockey, D.D., Heinzen, R.A., and Hackstadt, T. (1995) Cloning and characterization of a *Chlamydia psittaci* gene coding for a protein localized in the inclusion membrane of infected cells. *Mol Microbiol* **15**: 617–626.

Rosario, C.J., and Tan, M. (2012) The early gene product EUO is a transcriptional repressor that selectively regulates promoters of *Chlamydia* late genes. *Mol Microbiol* **84**: 1097–1107.

Seth-Smith, H.M.B., Harris, S.R., Persson, K., Marsh, P., Barron, A., Bignell, A., *et al.* (2009) Co-evolution of genomes and plasmids within *Chlamydia trachomatis* and the emergence in Sweden of a new variant strain. *BMC Genomics* **10**: 239.

Shaw, A.C., Christiansen, G., Roepstorff, P., and Birkelund, S. (2000) Genetic differences in the *Chlamydia trachomatis* tryptophan synthase alfa-subunit can explain variations in serovar pathogenesis. *Microbes Infect* **2**: 581–592.

Shaw, E.I., Dooley, C.A., Fischer, E.R., Scidmore, M.A., Fields, K.A., and Hackstadt, T. (2000) Three temporal classes of gene expression during the *Chlamydia trachomatis* developmental cycle. *Mol Microbiol* **37**: 913–925.

Somboonna, N., Wan, R., Ojcius, D.M., Pettengill, M. a, Joseph, S.J., Chang, A., *et al.* (2011) Hypervirulent *Chlamydia trachomatis* Clinical Strain Is a Recombinant between Lymphogranuloma Venereum (L2) and D Lineages. *MBio* **2**: 1–12.

Stephens, R.S., Kalman, S., Lammel, C., Fan, J., Marathe, R., Aravind, L., *et al.* (1998) Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*. *Science* **282**: 754–759.

Subtil, A., Delevoye, C., Balañá, M.E., Tastevin, L., Perrinet, S., and Dautry-Varsat, A. (2005) A directed screen for chlamydial proteins secreted by a type III mechanism identifies a translocated protein and numerous other new candidates. *Mol Microbiol* **56**: 1636–1647.

Subtil, A., Parsot, C., and Dautry-Varsat, A. (2001) Secretion of predicted Inc proteins of *Chlamydia pneumoniae* by a heterologous type III machinery. *Mol Microbiol* **39**: 792–800.

Thomson, N.R., Holden, M.T.G., Carder, C., Lennard, N., Lockey, S.J., Marsh, P., *et al.* (2008) *Chlamydia trachomatis*: genome sequence analysis of lymphogranuloma venereum isolates. *Genome Res* **18**: 161–171.

Unemo, M., Seth-Smith, H.M.B., Cutcliffe, L.T., Skilton, R.J., Barlow, D., Goulding, D., *et al.* (2010) The Swedish new variant of *Chlamydia trachomatis*: Genome sequence, morphology, cell tropism and phenotypic characterization. *Microbiology* **156**: 1394–1404.

Weber, M.M., Bauler, L.D., Lam, J., and Hackstadt, T. (2015) Expression and localization of predicted inclusion membrane proteins in *Chlamydia trachomatis*. *Infect Immun* IAI.01075–15.

Annexes

Table A.1 – List of primers used in this work.

Lab #	Description	Chapter	Sequence (5'-3')[a]
7	Yscu_Fw; upstream from *yscU* gene, for sequencing.	2	AAAAAGCGCGATAGCAAGCC
8	Yscu_Rv; downstream from *yscU* gene, for sequencing.	2	ATAAGTGAACCTCTTGTTGG
22	IncA_FL_Fw; used to construct pFA1 with a NdeI restriction site	2	GGAATTCCATATGACAACGCCTACTCTAATCG
23	IncA_FL_Rv; used to construct pFA1 with a HindIII restriction site	2	CCCAAGCTTAGTAGACTAGTTCTGCAGCC
24	IncC_FL_Fw; used to construct pFA2 with a NdeI restriction site	2	GGAATTCCATATGACGTACTCTATGTCCG
25	IncC_FL_Rv; used to construct pFA2 with a HindIII restriction site	2	CCCAAGCTTTTACAGAAAAGAAGCATGC
26	TEM1_Fw; used to construct pFA3 with a NdeI restriction site	2	GGAATTCCATATGCACCCAGAAACGCTGGTG
28	pLJM3_Fw; upstream from multiple cloning site of pLJM3, for cloning and sequencing	2	GATTAAGTTGGGTAACGCC
29	pLJM3_Rv; downstream from multiple cloning site of pLJM3, for cloning and sequencing	2	TTGTGTGGAATTGTGAGCG
30	IncA10_Rv; used to construct pFA8	2	CACCAGCGTTTCTGGGTGAGGAGTCACGATTAGAGTAGG
31	TEM1_IncA10_Fw; used to construct pFA8	2	CCTACTCTAATCGTGACTCCTCACCCAGAAACGCTGGTG
32	IncA20_Rv; used to construct pFA9	2	CACCAGCGTTTCTGGGTGGGCTGAGTAGGAAGGTGCAGG
33	TEM1_IncA20_Fw; used to construct pFA9	2	CCTGCACCTTCCTACTCAGCCCACCCAGAAACGCTGGTG
34	IncA40_Rv; used to construct pFA10	2	CACCAGCGTTTCTGGGTGGGAGGCAATGGCTGCTATTTTC
35	TEM1_IncA40_Fw; used to construct pFA10	2	GAAAATAGCAGCCATTGCCTCCCACCCAGAAACGCTGGTG
36	IncC10_Rv; used to construct pFA11	2	CACCAGCGTTTCTGGGTGGTGTGCTATATCGGACATAGAGTACG
37	TEM1_IncC10_Fw; used to construct pFA11	2	CGTACTCTATGTCCGATATAGCACACCACCCAGAAACGCTGGTG
38	IncC20_Rv; used to construct pFA12	2	CACCAGCGTTTCTGGGTGGGGAGACGTGGGATTAGAAATATC
39	TEM1_IncC20_Fw; used to construct pFA12	2	GATATTTCTAATCCCACGTCTCCCCACCCAGAAACGCTGGTG
40	IncC40_Rv; used to construct pFA13	2	CACCAGCGTTTCTGGGTGTAAAGAGCCCACGGCAGAAGG
41	TEM1_IncC40_Fw; used to construct pFA13	2	CCTTCTGCCGTGGGCTCTTTACACCCAGAAACGCTGGTG
42	YopE15_Rv; used to construct pFA14	2	CACCAGCGTTTCTGGGTGTGCCGGC

Lab #	Description	Chapter	Sequence (5'-3')[a]
			AGGGGCAGTGATGTAG
43	TEM1_YopE15_Fw; used to construct pFA14	2	CTACATCACTGCCCCTGCCGGCACA CCCAGAAACGCTGGTG
60	TEM1_Rv_HindIII; used to construct pFA03 and TEM-1 hybrids with a HindIII restriction site	2	CCCAAGCTTTTACCAATGCTTAATCA GTGAGG
66	SycT-20-TEM1_Fw_NdeI; used to construct pFA19 with a NdeI restriction site	2	GAATTCCATATGCAGACAACCTTCAC AGAACTTATGCAACAGCTTTTCCTGA AGCTTGGCTTGAACCATCACCCAGA AACGCTGGTG
67	TEM1_Rv_XhoI; used to construct TEM-1 hybrids with a XhoI restriction site	2	GATCCTCGAGTTACCAATGCTTAATC AGTGAGG
94	TEM1_Rv_Seq; complementary to *tem-1* sequence, for sequencing	2	CGATCAAGGCGAGTTACAT
102	RplJ-20-TEM1_Fw_NdeI; used to construct pFA37 with a NdeI restriction site	2	GGAATTCCATATGAAAGAAGAGAAA AAGTTGCTGCTTCGCGAGGTTGAAG AAAAGATAACCGCTTCTCAACACCCA GAAACGCTGGTG
429	CT005-20-TEM1_Fw_NdeI; used to construct pFA75 with a NdeI restriction site	2	GGAATTCCATATGACTCCAGTAACAC CAGTCCCTCCCCAATCTCCCCAACA GGTAAAAGGGCTTTTATCCCACCCA GAAACGCTGGTG
430	CT006-20-TEM1_Fw_NdeI; used to construct pFA76 with a NdeI restriction site	2	GGAATTCCATATGCCCTCCACTGTT GCACCTATAAAAGGACAGGATCACT TTTTAAATTTAGTTTTTCCTCACCCAG AAACGCTGGTG
431	CT036-20-TEM1_Fw_NdeI; used to construct pFA77 with a NdeI restriction site	2	GGAATTCCATATGTTCCCAATAGAAT GTCACACCCTACAGACTTCTTTTAAA CAAGTTCTCTCTTTGGTTCACCCAGA AACGCTGGTG
432	CT101-20-TEM1_Fw_NdeI; used to construct pFA78 with a NdeI restriction site	2	GGAATTCCATATGATCTCCATGATTC CAAGGTTTAGCATTGCCTGCATCCC CTTAGCTGTGTGGTTGTTTCACCCA GAAACGCTGGTG
433	CT134-20-TEM1_Fw_NdeI; used to construct pFA79 with a NdeI restriction site	2	GGAATTCCATATGGCTTGTTGCGCA TGTGTATACGGTTACGATGAGATTTG TTGTAGAGAGGAGACTGCACACCCA GAAACGCTGGTG
434	CT135-20-TEM1_Fw_NdeI; used to construct pFA80 with a NdeI restriction site	2	GGAATTCCATATGGTAAGCTTCGATT TAAATGATCCAGTAAGAAATACAGAC AATCATTACAGAAATATCCACCCAGA AACGCTGGTG
435	CT164-20-TEM1_Fw_NdeI; used to construct pFA81 with a NdeI restriction site	2	GGAATTCCATATGACGATATCCTCCA CAACAACACCTTCTAAAAGTAGTTCT TGTTTTCAGTTTAGAGAACACCCAGA AACGCTGGTG
436	CT179-20-TEM1_Fw_NdeI; used to	2	GGAATTCCATATGCTTGCAGCGGTA

Lab #	Description	Chapter	Sequence (5'-3')[a]
	construct pFA82 with a NdeI restriction site		GGAGCTTTTTTAGCTTTGTGCATAGG TGGTGTACTGTGCTGTATTCACCCA GAAACGCTGGTG
437	CT195-20-TEM1_Fw_NdeI; used to construct pFA83 with a NdeI restriction site	2	GGAATTCCATATGGTTTCTATGTCTC TAAATCTCCCTCCTGCGGAGGTTAG ATTGCGTCCCGTGACCGCGCACCCA GAAACGCTGGTG
438	CT196-20-TEM1_Fw_NdeI; used to construct pFA84 with a NdeI restriction site	2	GGAATTCCATATGGCAACAACAGTTA ATCCTAATTACTCTCTTCCTTTTTGTG AGAAAATGGTTTCCTCACACCCAGA AACGCTGGTG
439	CT214-20-TEM1_Fw_NdeI; used to construct pFA85 with a NdeI restriction site	2	GGAATTCCATATGCGAACAGACTCT CCTCTCAATCCTCCCGACTCTACTAG AGGAGTTTTTCAGTTTTTACACCCAG AAACGCTGGTG
440	CT222-20-TEM1_Fw_NdeI; used to construct pFA86 with a NdeI restriction site	2	GGAATTCCATATGCGTTGCTGTTGT GTTCGTACAAATTGTGAAGAAGTAAG ATCGTCGAGTACGGGAGATCACCCA GAAACGCTGGTG
441	CT224-20-TEM1_Fw_NdeI; used to construct pFA87 with a NdeI restriction site	2	GGAATTCCATATGAGTTTTGTTGGAG ATAGTGTACCTCTTAGAAGTTATATG CCAGAGGCTCCTTTGGTGCACCCAG AAACGCTGGTG
442	CT227-20-TEM1_Fw_NdeI; used to construct pFA88 with a NdeI restriction site	2	GGAATTCCATATGTCTTATCTTTTTT GTTCCTCTTGCGCTCCAACTCTGGA GAGCCCCGCAGAGTTGTGTCACCCA GAAACGCTGGTG
443	CT300-20-TEM1_Fw_NdeI; used to construct pFA89 with a NdeI restriction site	2	GGAATTCCATATGTGTTATGTCCTCA ATTTCATCACCGGGAAATATTCAACA CCCTCTCCCAACAACTGCCACCCAG AAACGCTGGTG
444	CT345-20-TEM1_Fw_NdeI; used to construct pFA90 with a NdeI restriction site	2	GGAATTCCATATGCAACTTCCGTCTA TTATTCAGTCTTTCTTCTTCCCTAAA GCTCCCCCATCTCCACTTCACCCAG AAACGCTGGTG
445	CT357-20-TEM1_Fw_NdeI; used to construct pFA91 with a NdeI restriction site	2	GGAATTCCATATGCCTGTAGTACAG AAACCTTCAGTTTTGGAGTACGCTCC TGTTTCTCCTTCTACGACTCACCCAG AAACGCTGGTG
446	CT383-20-TEM1_Fw_NdeI; used to construct pFA92 with a NdeI restriction site	2	GGAATTCCATATGTTCGGATCTATCC CTTGTTATCCCGGATACAACAATATT CCCGCATACAGTAACAGTCACCCAG AAACGCTGGTG
447	CT449-20-TEM1_Fw_NdeI; used to construct pFA93 with a NdeI restriction site	2	GGAATTCCATATGAAATTACCAGAAG TGAGTTTTAGTTTGCCTACAGCTGTT TGGGCATCTTCAACAAAACACCCAG AAACGCTGGTG

Lab #	Description	Chapter	Sequence (5'-3')[a]
448	CT483-20-TEM1_Fw_NdeI; used to construct pFA94 with a NdeI restriction site	2	GGAATTCCATATGGATTTTATGTCTG TTGTCCCACAGAGTCCTTGTTCATCT CCTACAAACTTCTATAGGCACCCAG AAACGCTGGTG
449	CT484-20-TEM1_Fw_NdeI; used to construct pFA95 with a NdeI restriction site	2	GGAATTCCATATGAAATCACGAAAC GCTCAGTCCATATTAGAGTCTTTATG TAAGAAAACGCATCGACTCCACCCA GAAACGCTGGTG
450	CT565-20-TEM1_Fw_NdeI; used to construct pFA96 with a NdeI restriction site	2	GGAATTCCATATGGCGGACGAGATG CAAAAAGAGAGCTCCTCACAAGAAC CATCGGCCTCTAAGTTTGGTCACCC AGAAACGCTGGTG
451	CT728-20-TEM1_Fw_NdeI; used to construct pFA97 with a NdeI restriction site	2	GGAATTCCATATGGCATCAGAATAC GAGCACTTCGGTAACTTATCTCCGG AGGATCATGTCAAAGAAGTTCACCC AGAAACGCTGGTG
452	CT789-20-TEM1_Fw_NdeI; used to construct pFA98 with a NdeI restriction site	2	GGAATTCCATATGAATTCAAATATAG AATATAGGCAATATCGTATAGATATA CTGAGCTGTTTTATCTGCCACCCAG AAACGCTGGTG
453	CT850-20-TEM1_Fw_NdeI; used to construct pFA99 with a NdeI restriction site	2	GGAATTCCATATGGGATTCGGAACT GTGAGAGGGAAAGGAAAAGCTGTGA AGTCCTTCTTCCTAAGGCCTCACCC AGAAACGCTGGTG
515	CT058-20-TEM1_Fw_NdeI; used to construct pFA112 with a NdeI restriction site	2	GGAATTCCATATGTTTACATCGCTGT CCGCGATACAGAATGCTATACGTCC TTCTTGTCAACTTCCTGTTCACCCAG AAACGCTGGTG
516	CT117-20-TEM1_Fw_NdeI; used to construct pFA113 with a NdeI restriction site	2	GGAATTCCATATGGGAGACGTGATG ATACAGAGCGTGAAAACAGAAAGTG GGTTAGTTGATGGTCATCACCACCC AGAAACGCTGGTG
517	CT225-20-TEM1_Fw_NdeI; used to construct pFA114 with a NdeI restriction site	2	GGAATTCCATATGGCTAACAACTCCT TTATTCACCGTTCTAAAACATACCAG TTATTTGTAGTTGTTCTACACCCAGA AACGCTGGTG
518	CT226-20-TEM1_Fw_NdeI; used to construct pFA115 with a NdeI restriction site	2	GGAATTCCATATGCGAAATAGAGGC GCCATGTTTAATATTTCTTTTTGTTGT AATTCCAGCAAACCATTGCACCCAG AAACGCTGGTG
519	CT228-20-TEM1_Fw_NdeI; used to construct pFA116 with a NdeI restriction site	2	GGAATTCCATATGAGTACTACTATTA GCGGAGACGCTTCTTCTTTACCGTT GCCAACAGCTTCCTGCGTACACCCA GAAACGCTGGTG
520	CT232-20-TEM1_Fw_NdeI; used to construct pFA117 with a NdeI restriction site	2	GGAATTCCATATGGTTCATTCTGTAT ACAATTCATTGGCTCCAGAAGGTTTT AGCCAAGTCTCTATTCAACACCCAG

Lab #	Description	Chapter	Sequence (5'-3')[a]
			AAACGCTGGTG
521	CT249-20-TEM1_Fw_NdeI; used to construct pFA118 with a NdeI restriction site	2	GGAATTCCATATGGGTATCAAACCTC ATGATTATGGCTGTTGGGGAAGTAG AGGGAACGTCTTTACCTTACACCCA GAAACGCTGGTG
522	CT358-20-TEM1_Fw_NdeI; used to construct pFA119 with a NdeI restriction site	2	GGAATTCCATATGGCTACACCGATTA CTGTACCGCCTTCTTCAGCGTCATC ACAATCTTCTCCGGACGTTCACCCA GAAACGCTGGTG
523	CT365-20-TEM1_Fw_NdeI; used to construct pFA120 with a NdeI restriction site	2	GGAATTCCATATGGTTTCTAGGGTTC CTGGAAGTAGTGATCCTGTTGATGT AAATGCTCTTAGTAACCTTCACCCAG AAACGCTGGTG
524	CT440-20-TEM1_Fw_NdeI; used to construct pFA121 with a NdeI restriction site	2	GGAATTCCATATGAAAGTTGTTGTGA ATCCTACTCAAGAATATTCACAAATT TCAACTCCGGTATTGCCGCACCCAG AAACGCTGGTG
525	CT618-20-TEM1_Fw_NdeI; used to construct pFA122 with a NdeI restriction site	2	GGAATTCCATATGGCAGCAACGGTA CCCATAGCAAGCCCAGTGGGCAGAC TACTATCTTCCGCAACCGCACACCC AGAAACGCTGGTG
526	CT813-20-TEM1_Fw_NdeI; used to construct pFA123 with a NdeI restriction site	2	GGAATTCCATATGACTACTCTTCCCA ATACTTGTACTTCAAACTCCAATTCT ATAAATACTTTCACGAAACACCCAGA AACGCTGGTG
529	IncCD-20-TEM1_Fw_NdeI; used to construct pFA126 with a NdeI restriction site	2	GGAATTCCATATGACTAAGGTTTATG CGAATAGCATTCAGCAAGAGAGAGT TGTGGATAGGATAGCTCTTCACCCA GAAACGCTGGTG
531	CT192L2-20-TEM1_Fw_NdeI; used to construct pFA128 with a NdeI restriction site	2	GGAATTCCATATGCAATCGGTTGGA CAAGAAGCTTCACAGAATACATTGTC TTGGACAATTCGACCGCGACACCCA GAAACGCTGGTG
	CT005-A_Fw; used for RT-qPCR	3	ATCGCGAGCAATGGAAACAG
	CT005-B_Rr; used for RT-qPCR	3	CAGCTAACTCAACTTGTTCGAGCTT
	CT006-A_Fw; used for RT-qPCR	3	TCAGGGTTGCTCCGGATATG
	CT006-B_Rv; used for RT-qPCR	3	AAGGAGAAAAGGGACAGAAAGGTT
	CT036-A_Fw; used for RT-qPCR	3	CGCATCGTGTGGATTCCTT
	CT036-B_Rv; used for RT-qPCR	3	GCAAAGTCCGGAGGCTATGA
	CT058-A_Fw; used for RT-qPCR	3	CCTGCTGCGATTGCAAATG
	CT058-B_Rv; used for RT-qPCR	3	CCTAATCCTCCTTGGCCTCTCT
	CT101-A_Fw; used for RT-qPCR	3	ACCGCTGCTATCTTTTCTGATGT
	CT101-B_Rv; used for RT-qPCR	3	TTGAACAGAGCGTAGCTAACAGAAG
	CT115-A_Fw; used for RT-qPCR	3	AGAGCGGTTGCATCGATCAC
	CT115-B_Rv; used for RT-qPCR	3	TTACAGCATAAGTTGTTCCACATCAA
	CT116-A_Fw; used for RT-qPCR	3	GTTTGGTGGGATGTCCATTTAAG
	CT116-B_Rv; used for RT-qPCR	3	CATTAGCTGGACTCGATTTGCAT

Lab #	Description	Chapter	Sequence (5'-3')[a]
	CT117-A_Fw; used for RT-qPCR	3	TGTAGGGCCTTTAGCCGTTTTA
	CT117-B_Rv; used for RT-qPCR	3	CACAAGAAGCTCGCCAGTGTT
	CT118-A_Fw; used for RT-qPCR	3	GCGTACAATCAATGCCTGTTATAGA
	CT118-B_Rv; used for RT-qPCR	3	GCCGTTTGTAAGCATTTCGTTA
	CT119-A_Fw; used for RT-qPCR	3	GCAGCCATTGCCTCCCTAA
	CT119-B_Rv; used for RT-qPCR	3	GCCAACAAGATGTCCCAAAAG
	CT134.A_Fw; used for RT-qPCR	3	CGGTTACGATGAGATTTGTTGTAGA
	CT134-B_Rv; used for RT-qPCR	3	CGCCACTACTTTTTCTGCAGTCT
	CT135-A_Fw; used for RT-qPCR	3	ACGAACGGATCATGTTTGAAGA
	CT135-B_Rv; used for RT-qPCR	3	CGGCTTCGAGAACACTAGGAACT
	CT164-C_Fw; used for RT-qPCR	3	TGTGAAGACGTATCGGGTTACATC
	CT164-D_Rv; used for RT-qPCR	3	CAATAGGAGCAAGGAGCTAATCATT
	CT179-A_Fw; used for RT-qPCR	3	CCCATATTCGAGAAGGGCTCTA
	CT179-B_Rv; used for RT-qPCR	3	CGTGATCTCGTTGTTCCGATT
	CT192-A_Fw; used for RT-qPCR	3	TGGGCAGCCATGTAGTTGAA
	CT192-B_Rv; used for RT-qPCR	3	CATTACGGCCTATTTTACCAGCTTT
	CT195-A_Fw; used for RT-qPCR	3	CAGTTTGGAACGAGATTGGAATAA
	CT195-B_Rv; used for RT-qPCR	3	CAAACTCCTCGCGAACGAAT
	CT196-A_Fw; used for RT-qPCR	3	CCTCTATTGCAGCAAGCTACTTTCT
	CT196-B_Rv; used for RT-qPCR	3	CGATAACCAAAGTAACTCCCAGAGA
	CT214-A_Fw; used for RT-qPCR	3	TTATTTCCGGACAAGCAGATGA
	CT214-B_Rv; used for RT-qPCR	3	TTTTGATCCCAATCCGATTAGG
	CT222-A_Fw; used for RT-qPCR	3	GATTGGAGCACTTCTATTGGGAAT
	CT222-B_Rv; used for RT-qPCR	3	CCTATCAAACCTGCAGAACTTCCT
	CT223-A_Fw; used for RT-qPCR	3	CCGCCTCTCTTCTTCTATGCA
	CT223-B_Rv; used for RT-qPCR	3	AACAGCTCCCAAAGCAACCA
	CT224-A_Fw; used for RT-qPCR	3	CCTTTGGTGGATAGCGCTTCT
	CT224-B_Rv; used for RT-qPCR	3	GAGCGCAAGCACAGCAATAC
	CT225-A_Fw; used for RT-qPCR	3	CGCAATTCAGATCAGCTATTACAAG
	CT225-B_Rv; used for RT-qPCR	3	AATTTTCCGCGCACTTTCC
	CT226-A_Fw; used for RT-qPCR	3	AACCGACGAGACTCCCTTCA
	CT226-B_Rv; used for RT-qPCR	3	TCACGGCCTCTGGAGCTAAC
	CT227-A_Fw; used for RT-qPCR	3	GAGAGCCCCGCAGAGTTGT
	CT227-B_Rv; used for RT-qPCR	3	TCGTTTGCAATAAATATGCGTCTT
	CT228-A_Fw; used for RT-qPCR	3	TCGTAGGCGCTTTAGTTGTTGTC
	CT228-B_Rv; used for RT-qPCR	3	CGCACAAAACCAAAGCTAATACC
	CT229-A_Fw; used for RT-qPCR	3	AATCGGGACCCCAGCTTCT
	CT229-B_Rv; used for RT-qPCR	3	AGGGAACCAACCATAACTAAGATCA
	CT232-A_Fw; used for RT-qPCR	3	CAGCAATAGTCCTTTCCATCGTT
	CT232-B_Rv; used for RT-qPCR	3	CGGTGTTAAGTGCAGCAAGAAG
	CT233-A_Fw; used for RT-qPCR	3	TGGCTGGGTCAACTAATGCA
	CT233-B_Rv; used for RT-qPCR	3	TAGTGATGGAGGCGCTCACA
	CT249-A_Fw; used for RT-qPCR	3	GCCGCTCTCGCCTTTAAACT
	CT249-B_Rv; used for RT-qPCR	3	AGTAGGCACGTCGTGCTTGA
	CT288-A_Fw; used for RT-qPCR	3	GCCTGCCTTTTATCGCTGTTAT
	CT288-B_Rv; used for RT-qPCR	3	CCATCCCCAATGCTAAGGAA

Lab #	Description	Chapter	Sequence (5'-3')[a]
	CT300-A_Fw; used for RT-qPCR	3	TAACGCAGCAATCGCAACTG
	CT300-B_Rv; used for RT-qPCR	3	GCGGCAATGGTTATAAGACCTAAT
	CT345-A_Fw; used for RT-qPCR	3	TCTGGAGTATTCCTTTCCCTTAGC
	CT345-B_Rv; used for RT-qPCR	3	ACAAATGCCTGCGCCTAATAG
	CT357-A_Fw; used for RT-qPCR	3	GCCCTTGCTTTAGGATTATGGA
	CT357-B_Rv; used for RT-qPCR	3	CAGTAGTGCATGCTTGGTTGATG
	CT358-A_Fw; used for RT-qPCR	3	AATAGCGTTTCTGGCATCGAA
	CT358-B_Rv; used for RT-qPCR	3	GCAATAAGCAGGCTCCATCCTA
	CT365-A_Fw; used for RT-qPCR	3	AGTTAGTTGGGAAACCGGTGATT
	CT365-B_Rv; used for RT-qPCR	3	ACTCACGAATAGCTTGCTCTGTTTG
	CT383-A_Fw; used for RT-qPCR	3	GGCGTGGCAATAGCCTTCTA
	CT383-B_Rv; used for RT-qPCR	3	AGAGATCCGAGTGCACAAGAAAG
	CT440-E_Fw; used for RT-qPCR	3	CTAACCTGCAGCCTTTTGTTCA
	CT440-F_Rv; used for RT-qPCR	3	AATAATAAACAGGACTGCCGAACAC
	CT442-A_Fw; used for RT-qPCR	3	TCGAATCTTCTATCTTCCACTGCAT
	CT442-B_Rv; used for RT-qPCR	3	TCCTATCACCACTAATCCCACCTTA
	CT449-A_Fw; used for RT-qPCR	3	TTCAGTTCCTTGGATCGTGTACA
	CT449-B_Rv; used for RT-qPCR	3	CACTATTCGTTGCACCCAATTC
	CT483-A_Fw; used for RT-qPCR	3	TCTTGGCGAGCCTCTCTATCC
	CT483-B_Rv; used for RT-qPCR	3	AAGGCTCCCTACTTCGTATCGA
	CT484-A_Fw; used for RT-qPCR	3	ACTCTTATTGCGCTATCTGCTCAA
	CT484-B_Rv; used for RT-qPCR	3	CCAGCTCTGCGACCATTAAAGT
	CT565-A_Fw; used for RT-qPCR	3	CTATTACGCTAAAAACGGGCCTAT
	CT565-B_Rv; used for RT-qPCR	3	GCAACCCCGCATACCAAAG
	CT628-A_Fw; used for RT-qPCR	3	TTCCGCAACCGCAACTACTT
	CT618-B_Rv; used for RT-qPCR	3	GCTTTGAGCGAAGGGATGTT
	CT728-A_Fw; used for RT-qPCR	3	GAGCCGCTGATCTCACAAATG
	CT728-B_Rv; used for RT-qPCR	3	AAGAGAGGAGTCGGAACACACAA
	CT789-A_Fw; used for RT-qPCR	3	CGTCCTGTCCGCTGGTTCT
	CT789-B_Rv; used for RT-qPCR	3	TCCAGAGGCAATGCCAAAG
	CT813-A_Fw; used for RT-qPCR	3	ATCCAAGGCTCTCGGAAGGA
	CT813-B_Rv; used for RT-qPCR	3	TCCTGCGATCCATGAGGTTT
	CT850-A_Fw; used for RT-qPCR	3	TGACGGCATCGCACAATT
	CT850-B_Rv; used for RT-qPCR	3	CTTGCACTTCGCGGTTATAGG
	16SRNA-9_Fw; used for RT-qPCR	3	GCGAAGGCGCTTTTCTAATTTAT
	16SRNA-10_Rv; used for RT-qPCR	3	CCAGGGTATCTAATCCTGTTTGCT
1005	CT058_Fw; used for the transcriptional linkage; amplification of *ct058* from cDNA or genomic DNA	3	GAAATGCTTTTTGGATGG
1012	CT058_Rv; used for the transcriptional linkage; amplification of *ct058* from cDNA or genomic DNA	3	GAATGCATCGTTCACTCC
1004	CT059_CT058_Fw; used for the transcriptional linkage; amplification of *ct059-ct058* from cDNA or genomic DNA	3	GCGTGATTGAGGTCTTGG
1010	CT059_CT058_Rv; used for the	3	GACACCTAAAATAACAGC

Lab #	Description	Chapter	Sequence (5'-3')[a]
	transcriptional linkage; amplification of *ct059-ct058* from cDNA or genomic DNA; used for the transcription start site; amplification and sequencing of: *ct059* promoter region and coding sequence; *ct059-ct058* intergenic region; and *ct058* first 64 codons. In strains B/Har36, C/TW3, F/CS465-95, and L2b/CS019-08		
1011	CT059_SP1_Rv; used for the transcription start site; generation of single stranded (ss) cDNA from RNA	3	AGAATCTTCAGGATCACC
1003	CT059_SP2_Rv; used for the transcription start site; generation of double stranded (ds) cDNA from ss cDNA	3	CCAAGACCTCAATCACGC
1002	CT059_Prom_Fw; used for the transcription start site; amplification and sequencing of: *ct059* promoter region and coding sequence; *ct059-ct058* intergenic region; and *ct058* first 64 codons. In strains B/Har36, C/TW3, F/CS465-95, and L2b/CS019-08	3	ATCAATCAGCCATCTAGG
852	CT192_Fw; used for the transcriptional linkage; amplification of *ct192* from cDNA or genomic DNA	3	CGCGGATCCCACATCCGAAGCAGCAGGG
948	CT192_Rv; used for the transcriptional linkage; amplification of *ct192* from cDNA or genomic DNA	3	CGCGGATCCTTAACAATCATTGGAAAC
1023	CT193_CT192_Fw; used for the transcriptional linkage; amplification of *ct193-ct192* from cDNA or genomic DNA; used for the transcription start site; amplification and sequencing of the region upstream from start codon of *ct192* in strains B/Har36, C/TW3, and L2b/CS019-08	3	AGAGATTAAAAGTTGCGG
1020	CT193_CT192_Rv; used for the transcriptional linkage; amplification of *ct193-ct192* from cDNA or genomic DNA; used for the transcription start site; amplification and sequencing of the region upstream from start codon of *ct192* in strains B/Har36, C/TW3, and L2b/CS019-08	3	GCTCCACCAACAGCTGCAGC
1034	CT192_SP1_Rv; used for the transcription start site; generation of single stranded (ss) cDNA from RNA	3	TTTCTGCAGAATGACTGG
1047	CT192_SP2_Rv; used for the transcription	3	TTGTTGGACTTCCTCCGG

Lab #	Description	Chapter	Sequence (5'-3')[a]
	start site; generation of double stranded (ds) cDNA from ss cDNA; amplification and sequencing of the region upstream from start codon of *ct192* in strain F/CS465-95		
1019	CT192_Prom2_Fw; used for the transcription start site; amplification and sequencing of the region upstream from start codon of *ct192* in strain F/CS465-95	3	AGGGATAGGAGATTTGCC
950	CT214_SP1_Rv; used for the transcription start site; generation of single stranded (ss) cDNA from RNA	3	CGCGGATCCTTAACCAAATAATGCAGG
1036	CT214_SP2_Rv; used for the transcription start site; generation of double stranded (ds) cDNA from ss cDNA; amplification and sequencing of the region upstream from start codon of *ct192* in strains B/Har36, C/TW3, F/CS465-95, and L2b/CS019-08	3	TACAGCAGAGACTAACCC
1035	CT214_Prom_Fw; used for the transcription start site; amplification and sequencing of the region upstream from start codon of *ct192* in strains B/Har36, C/TW3, F/CS465-95, and L2b/CS019-08	3	CTCCGTCAGAGCGGATGC
1483	2xHA-IncDTerm_Rv_Sal; used to construct pSVP247 with a SalI restriction site	4	GATCGTCGACGTCTTAGGAGCTTTTTGCAATGC
1486	2XHA-IncDTerm_Fwd_Not; used to construct pSVP247 with a NotI restriction site	4	GATCGCGGCCGCTATCCGTATGATGTGCCGGA
1487	2XHA-IncDTerm_Fwd_OL; used to construct pSVP247	4	GATGTTCCTGATTATGCTTAAGGATGACATGTGATTCGCGTAGG
1488	2XHA-IncDTerm_Rv_OL; used to construct pSVP247	4	CCTACGCGAATCACATGTCATCCTTAAGCATAATCAGG
1546	PromIncD_KpnI_Fwd; used to construct pSVP255 with a KpnI restriction site	4	GATCGGTACCAACGGAGCCTTCTAGCTATTTTG
1565	PromIncD-CT288_OL_Fw; used to construct pSVP255	4	ATCTGTCGAAGTGAGGTTTATGGTTTATTTTAGAGCTCATC
1566	PromIncD-CT288_OL_Rv; used to construct pSVP255	4	GATGAGCTCTAAAATAAACCATAAACCTCACTTCGACAGAT
1567	CT288_NotI_Rv; used to construct pSVP255 with a NotI restriction site	4	GATCGCGGCCGCGGTGATTATCTAACAGGTATTG
599	CT288_Seq1; complementary to *ct288* sequence, for sequencing	4	AGTTCTTGGTTCCAGATACG
600	CT288_Seq2; complementary to *ct288* sequence, for sequencing	4	TTCTGAGTCTAGAAGCTGCG
627	EGFP-C1_Fw; complementary to pEGFP-	4	TGGTCCTGCTGGAGTTCGTG

Lab #	Description	Chapter	Sequence (5'-3')[a]
	C1 sequence, for sequencing		
628	EGFP-C1_Rv; complementary to pEGFP-C1 sequence, for sequencing	4	TTATGTTTCAGGTTCAGGG
642	T7_Fw; complementary to pEF6/Myc-His C sequence, for sequencing	4	TAATACGACTCACTATAGGGC
644	pGADT7_Fw; complementary to pGADT7 sequence, for sequencing	4	CTATTCGATGATGAAGATACCCCAC CAAACC
645	pGADT7_RV; complementary to pGADT7 sequence, for sequencing	4	AGTGAACTTGCGGGGTTTTTCAGTAT CTACGA
862	CT288_del_241_292_Fw; used to construct pFA147, pFA184, pFA185, pFA139, and pFA197	4	GCCAGCCGCACACTATATAAGTATTT GCTCGAACATTCTCC
863	CT288_del_241_292_Rv; used to construct pFA147, pFA184, pFA185, pFA139, and pFA197	4	GGAGAATGTTCGAGCAAATACTTATA TAGTGTGCGGCTGGC
943	CT249_NdeI_Fw; used to construct pFA159 with a NdeI restriction site	4	GGAATTCCATATGGGTATCAAACCTC ATGATTATGGC
944	CT288_NdeI_Fw; used to construct pFA147, pFA178, pFA184, and pFA185 with a NdeI restriction site	4	GGAATTCCATATGAAAAAGGCTCTG GCTCAACG
952	CT249_BamHI_Rv; used to construct pFA159 with a BamHI restriction site	4	CGCGGATCCCTACGCCTTTGAAGTA ATCG
953	CT288_KpnI_Fw; used to construct pFA139 with a KpnI restriction site	4	GATCGGTACCAAAAAGGCTCTGGCT CAACG
954	CT288_BamHI_Rv; used to construct pFA147, pFA179, pFA184, pFA185, and pFA139 with a BamHI restriction site	4	CGCGGATCCTTAGTGATTATCTAACA GG
977	pGBKT7_Rv; complementary to pGBKT7 sequence, for sequencing	4	AAATCATAAGAAATTCGC
1212	CT228_NdeI_Fw; used to construct pFA155 with a NdeI restriction site	4	GGAATTCCATATGAGTACTACTATTA GCGGA
1216	CT228_BamHI_Rv; used to construct pFA155 with a BamHI restriction site	4	CGCGGATCCCTAAGAAGCTTGGTTA GC
1221	CT228_del_38_86_Fw; used to construct pFA155	4	GGGAATACTTGTTCCAAAATTCGTCT TATGTATCGATCC
1222	CT228_del_38_86_Rv; used to construct pFA155	4	GGATCGATACATAAGACGAATTTTG GAACAAGTATTCCC
1229	CT249_del_51_97_Fw; used to construct pFA159	4	GCAGCAAAAGTTGCTCGCAAACTTT ATGACCTATTCAAGC
1230	CT249_del_51_97_Rv; used to construct pFA159	4	GCTTGAATAGGTCATAAAGTTTGCGA GCAACTTTTGCTGC
1350	GFPC_CC_XhoI_Fw; used to construct pFA164 with a XhoI restriction site	4	GATCCTCGAGctGAAGACAGTAGCAC AGACAC
1351	GFPC_CC_EcoRI_Rv; used to construct pFA164, and pFA196 with a EcoRI restriction site	4	GATCGAATTCTCAGATTTCAACTGGC TTTATAACAGG
1352	CC_HA_EcoRI_Fw; used to construct	4	GATCGAATTCACCATGGAAGACAGT

Lab #	Description	Chapter	Sequence (5'-3')[a]
	pFA167 with a EcoRI restriction site		AGCACAGACAC
1354	CCD_HA_EcoRI_Fw; used to construct pFA167 with a EcoRI restriction site	4	GATCGAATTCACCATGGAAATTTGTG TGACCCAG
1355	CC_HA_NotI_Rv; used to construct pFA167, and pFA168 with a NotI restriction site	4	ATAAGAATGCGGCCGCTTAAGCATA ATCAGGAACATCATACGGATAGATTT CAACTGGCTTTATAACAGG
1356	CC_Seq1_Fw; complementary to *CCDC146* sequence, for sequencing	4	TAAGAAAGATGGAACTGC
1357	CC_Seq2_Rv; complementary to *CCDC146* sequence, for sequencing	4	TCTACTAACTTAGACTCC
1358	CC_Seq3_Rv; complementary to *CCDC146* sequence, for sequencing	4	ATGAAATCCTTCTCCAGC
1474	CT288_89-241_BamHI_Rv; used to construct pFA178 with a BamHI restriction site	4	CGCGGATCCTTAATATAGTGTGCGG CTGGC
1475	CT288_292-563_NdeI_Pf; used to construct pFA179 with a NdeI restriction site	4	GGAATTCCATATGAAGTATTTGCTCG AACATTCTCC
1546	PromIncD_KpnI_Fw; used to construct pSVP255 with a KpnI restriction site	4	GATCGGTACCAACGGAGCCTTCTAG CTATTTTG
1565	PromIncD-CT288_OL_Fw; used to construct pSVP255	4	ATCTGTCGAAGTGAGGTTTATGGTTT ATTTTAGAGCTCATC
1566	PromIncD-CT288_OL_Rv; used to construct pSVP255	4	GATGAGCTCTAAAATAAACCATAAAC CTCACTTCGACAGAT
1567	CT288_NotI_Rv; used to construct pSVP255 with a NotI restriction site	4	GATCGCGGCCGCGGTGATTATCTAA CAGGTATTG
1730	GFPC_C692_XhoI_Fw; used to construct pFA196 with a XhoI restriction site	4	GATCCTCGAGCTCAAATTTGTGTGA CCCAG
1733	CT288-2HA_BamHI_Fw; used to construct pFA197 with a BamHI restriction site	4	GATCGGATCCACCATGGCGAAGGCT CTGGCTCAAGC
1734	CT288-2HA_EcoRI_Rv; used to construct pFA197 with a EcoRI restriction site	4	GATCGAATTCTTAAGCATAATCAGGA ACATCATACGG

[a]Restriction sites are underlined; Kozak sequences are in bold; modifications to clone in frame are lower cased.

Table A.2 – List of plasmids used in this work.

Plasmid	Description/ Construction	Chapter	Reference
pLY16	*yscU* mutator. Construction of *Y. enterocolitica* ΔHOPEMT ΔYscU.	2	(Sorg *et al*., 2007)
pFA1001	*Y. enterocolitica* virulence plasmid that encodes for the T3S system of strain ΔHOPEMT, with a deletion of YscU codons 1-356	2	This work
pLJM3	Expresses YopE under the control of its own promoter (P_{yopE}).	2	(Marenne *et al*., 2003)
pCX340	Expresses mature TEM-1 β-lactamase.	2	(Charpentier and Oswald, 2004)
pFA1	Expresses IncA under the control of P_{yopE}. The *incA* gene was amplified by PCR from L2/434 chromosomal DNA using primers #22 and #23. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA2	Expresses IncC under the control of P_{yopE}. The *incC* gene was amplified by PCR from L2/434 chromosomal DNA using primers #24 and #25. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA3	Expresses mature TEM-1 β-lactamase under the control of P_{yopE}. The *tem*-1 gene was amplified by PCR from pCX340 DNA using primers #26 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA8	Expresses $IncA_{10}$-TEM-1 hybrid under the control of P_{yopE}. Nucleotides 1 to 30 of *incA* were amplified by PCR from pFA1 DNA using primers #28 and #30. Also *tem*-1 was amplified by PCR from pCX340 using primers #31 and #60. The DNA products were then fused by overlapping PCR using primers #28 and #60. The final DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA9	Expresses $IncA_{20}$-TEM-1 hybrid under the control of P_{yopE}. Nucleotides 1 to 60 of *incA* were amplified by PCR from pFA1 DNA using primers #28 and #32. Also *tem*-1 was amplified by PCR from	2	This work

Plasmid	Description/ Construction	Chapter	Reference
	pCX340 DNA using primers #33 and #60. The DNA products were then fused by overlapping PCR using primers #28 and #60. The final product was digested NdeI-HindIII and ligated into those sites of pLJM3.		
pFA10	Expresses $IncA_{40}$-TEM-1 hybrid under the control of P_{yopE}. Nucleotides 1 to 120 of *incA* were amplified by PCR from pFA1 DNA using primers #28 and #34. Also *tem*-1 was amplified by PCR from pCX340 using primers #35 and #60. The DNA products were then fused by overlapping PCR using primers #28 and #60. The final DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA11	Expresses $IncC_{10}$-TEM-1 hybrid under the control of P_{yopE}. Nucleotides 1 to 30 of *incC* were amplified by PCR from pFA2 DNA using primers #28 and #36. Also *tem*-1 was amplified by PCR from pCX340 using primers #37 and #60. The DNA products were then fused by overlapping PCR using primers #28 and #60. The final DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA12	Expresses $IncC_{20}$-TEM-1 hybrid under the control of P_{yopE}. Nucleotides 1 to 60 of *incC* were amplified by PCR from pFA2 DNA using primers #28 and #38. Also *tem*-1 was amplified by PCR from pCX340 using primers #39 and #60. The DNA products were then fused by overlapping PCR using primers #28 and #60. The final DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA13	Expresses $IncC_{40}$-TEM-1 hybrid under the control of P_{yopE}. Nucleotides 1 to 120 of *incC* were amplified by PCR from pFA2 DNA using primers #28 and #40. Also *tem*-1 was amplified by PCR from pCX340 using primers #41 and #60. The DNA products were then fused by overlapping PCR using primers #28 and #60. The final DNA product was digested	2	This work

Plasmid	Description/ Construction	Chapter	Reference
	with NdeI-HindIII and ligated into those sites of pLJM3.		
pFA14	Expresses $YopE_{15}$-TEM-1 hybrid under the control of P_{yopE}. Nucleotides 1 to 45 of *yopE* were amplified by PCR from pLJM3 DNA using primers #28 and #42. Also *tem*-1 was amplified by PCR from pCX340 using primers #43 and #60. The DNA products were then fused by overlapping PCR using primers #28 and #60. The final DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA19	Expresses $SycT_{20}$-TEM-1 hybrid under the control of P_{yopE}. The *$sycT_{20}$-tem-1* was amplified by PCR from pCX340 using primers #66 and #67. The DNA product was digested with NdeI-XhoI and ligated into those sites of pLJM3.	2	This work
pFA37	Expresses $RplJ_{20}$-TEM-1 hybrid under the control of P_{yopE}. The *$rplJ_{20}$-tem-1* was amplified by PCR from pCX340 using primers #102 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA75	Expresses $CT005_{20}$-TEM-1 hybrid under the control of P_{yopE}. The *$ct005_{20}$-tem-1* was amplified by PCR from pCX340 using primers #429 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA76	Expresses $CT006_{20}$-TEM-1 hybrid under the control of P_{yopE}. The *$ct006_{20}$-tem-1* was amplified by PCR from pCX340 using primers #430 and #67. The DNA product was digested with NdeI-XhoI and ligated into those sites of pLJM3.	2	This work
pFA77	Expresses $CT036_{20}$-TEM-1 hybrid under the control of P_{yopE}. The *$ct036_{20}$-tem-1* was amplified by PCR from pCX340 using primers #431 and #67. The DNA product was digested with NdeI-XhoI and ligated into those sites of pLJM3.	2	This work
pFA78	Expresses $CT101_{20}$-TEM-1 hybrid under the control of P_{yopE}. The *$ct101_{20}$-tem-1* was amplified by PCR from pCX340 using primers #432 and #67. The DNA product	2	This work

Plasmid	Description/ Construction	Chapter	Reference
	was digested with NdeI-XhoI and ligated into those sites of pLJM3.		
pFA79	Expresses CT134$_{20}$-TEM-1 hybrid under the control of P$_{yopE}$. The *ct134$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #433 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA80	Expresses CT135$_{20}$-TEM-1 hybrid under the control of P$_{yopE}$. The *ct135$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #434 and #67. The DNA product was digested with NdeI-XhoI and ligated into those sites of pLJM3.	2	This work
pFA81	Expresses CT164$_{20}$-TEM-1 hybrid under the control of P$_{yopE}$. The *ct164$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #435 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA82	Expresses CT179$_{20}$-TEM-1 hybrid under the control of P$_{yopE}$. The *ct179$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #436 and #67. The DNA product was digested with NdeI-XhoI and ligated into those sites of pLJM3.	2	This work
pFA83	Expresses CT195$_{20}$-TEM-1 hybrid under the control of P$_{yopE}$. The *ct195$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #437 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA84	Expresses CT196$_{20}$-TEM-1 hybrid under the control of P$_{yopE}$. The *ct196$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #438 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA85	Expresses CT214$_{20}$-TEM-1 hybrid under the control of P$_{yopE}$. The *ct214$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #439 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA86	Expresses CT222$_{20}$-TEM-1 hybrid under the control of P$_{yopE}$. The *ct222$_{20}$-tem-1* was amplified by PCR from pCX340 using	2	This work

Plasmid	Description/ Construction	Chapter	Reference
	primers #440 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.		
pFA87	Expresses CT224$_{20}$-TEM-1 hybrid under the control of P$_{yopE}$. The *ct224$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #441 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA88	Expresses CT227$_{20}$-TEM-1 hybrid under the control of P$_{yopE}$. The *ct227$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #442 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA89	Expresses CT300$_{20}$-TEM-1 hybrid under the control of P$_{yopE}$. The *ct300$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #443 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA90	Expresses CT345$_{20}$-TEM-1 hybrid under the control of P$_{yopE}$. The *ct345$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #444 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA91	Expresses CT357$_{20}$-TEM-1 hybrid under the control of P$_{yopE}$. The *ct357$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #445 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA92	Expresses CT383$_{20}$-TEM-1 hybrid under the control of P$_{yopE}$. The *ct383$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #446 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA93	Expresses CT449$_{20}$-TEM-1 hybrid under the control of P$_{yopE}$. The *ct449$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #447 and #67. The DNA product was digested with NdeI-XhoI and ligated into those sites of pLJM3.	2	This work
pFA94	Expresses CT483$_{20}$-TEM-1 hybrid under the control of P$_{yopE}$. The *ct483$_{20}$-tem-1*	2	This work

Plasmid	Description/ Construction	Chapter	Reference
	was amplified by PCR from pCX340 using primers #448 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.		
pFA95	Expresses CT484$_{20}$-TEM-1 hybrid under the control of P$_{yopE}$. The *ct484$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #449 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA96	Expresses CT565$_{20}$-TEM-1 hybrid under the control of P$_{yopE}$. The *ct565$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #450 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA97	Expresses CT728$_{20}$-TEM-1 hybrid under the control of P$_{yopE}$. The *ct728$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #451 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA98	Expresses CT789$_{20}$-TEM-1 hybrid under the control of P$_{yopE}$. The *ct789$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #452 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA99	Expresses CT850$_{20}$-TEM-1 hybrid under the control of P$_{yopE}$. The *ct850$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #453 and #67. The DNA product was digested with NdeI-XhoI and ligated into those sites of pLJM3.	2	This work
pFA112	Expresses CT058$_{20}$-TEM-1 hybrid under the control of P$_{yopE}$. The *ct058$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #515 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA113	Expresses CT117$_{20}$-TEM-1 hybrid under the control of P$_{yopE}$. The *ct117$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #516 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA114	Expresses CT225$_{20}$-TEM-1 hybrid under	2	This work

Plasmid	Description/ Construction	Chapter	Reference
	the control of P_{yopE}. The *ct225$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #517 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.		
pFA115	Expresses CT226$_{20}$-TEM-1 hybrid under the control of P_{yopE}. The *ct226$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #518 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA116	Expresses CT228$_{20}$-TEM-1 hybrid under the control of P_{yopE}. The *ct228$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #519 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA117	Expresses CT232$_{20}$-TEM-1 hybrid under the control of P_{yopE}. The *ct232$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #520 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA118	Expresses CT249$_{20}$-TEM-1 hybrid under the control of P_{yopE}. The *ct249$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #521 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA119	Expresses CT358$_{20}$-TEM-1 hybrid under the control of P_{yopE}. The *ct358$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #522 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA120	Expresses CT365$_{20}$-TEM-1 hybrid under the control of P_{yopE}. The *ct365$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #523 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA121	Expresses CT440$_{20}$-TEM-1 hybrid under the control of P_{yopE}. The *ct440$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #524 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work

Plasmid	Description/ Construction	Chapter	Reference
pFA122	Expresses CT618$_{20}$-TEM-1 hybrid under the control of P$_{yopE}$. The *ct618$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #525 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA123	Expresses CT813$_{20}$-TEM-1 hybrid under the control of P$_{yopE}$. The *ct813$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #526 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA126	Expresses IncDD$_{20}$-TEM-1 hybrid under the control of P$_{yopE}$. The *incDD$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #529 and #60. The DNA product was digested with NdeI-HindIII and ligated into those sites of pLJM3.	2	This work
pFA128	Expresses CT192L2$_{20}$-TEM-1 hybrid under the control of P$_{yopE}$. The *ct192L2$_{20}$-tem-1* was amplified by PCR from pCX340 using primers #531 and #67. The DNA product was digested with NdeI-XhoI and ligated into those sites of pLJM3.	2	This work
pGADT7	Yeast-two hybrid (Y2H) vector for expression of Gal4AD-myc fusion proteins.	4	(Clontech)
pGBKT7	Y2H vector for expression of Gal4BD-HA fusion proteins.	4	(Clontech)
pEGFP-C1	Mammalian transfection vector for expression of EGFP fusion proteins, under the control of the *CMV* promoter.	4	(Clontech)
pEF6/Myc-His C	Mammalian transfection vector for expression of Myc-, His-, or HA-tagged proteins, under the control of the *EF-1α* promoter.	4	(Life Technologies)
pCCDC146	Contains the cDNA of *CCDC146*; clone IRATp970G0632D.	4	(Source BioScience)
p2TK2--SW2	*C. trachomatis-E. coli* shuttle vector encoding red shifted GFP (RSGFP) under the control of the promoter of *incD* (P$_{incD}$).	4	(Agaisse and Derré, 2013)
PSVP247	Modified *C. trachomatis-E. coli* shuttle vector for expression of HA-tagged proteins under the control of P$_{incD}$. Contains also the terminator sequence of *incD*. The DNA Sequence of 2xHA was amplified by PCR from pLJ1003 DNA	4	This Work

Plasmid	Description/ Construction	Chapter	Reference
	using primers #1486 e #1488. Also the DNA sequence of the *incD* terminator was amplified by PCR from L2/434 chromosomal DNA using primers #1487 and #1483. The DNA products were then fused by overlapping PCR using primers #1486 and #1483. The final DNA product was digested with NotI-SalI and ligated into those sites of p2TK2—SW2.		
pSVP255	*C. trachomatis-E. coli* shuttle vector expressing full-length CT288 with a C-terminal 2HA tag ($CT288_{FL}$-2HA) under the control of P_{incD}. The *ct288* gene and the P_{incD} were amplified by PCR from L2/434 chromosomal DNA using primers #1565 and #1567, and #1546 and #1566, respectively. The DNA products were then fused by overlapping PCR using primers #1546 and #1567. The final DNA product was digested with KpnI-NotI and ligated into those sites of pSVP247.	4	This Work
pFA147	Y2H vector expressing Gal4BD-myc-$CT288_{\Delta N\Delta TMD}$. Nucleotides 265 to 723, and 874 to 1689 of *ct288* were amplified by PCR from L2/434 chromosomal DNA using primers #944 and #863, and #862 and #954, respectively. The DNA products were then fused by overlapping PCR using primers #944 and #954. The final DNA product was digested with NdeI-BamHI and ligated into those sites of pGBKT7.	4	This Work
pFA155	Y2H vector expressing Gal4BD-myc-$CT228_{\Delta TMD}$. Nucleotides 1 to 111, and 259 to 1773 of *ct228* were amplified by PCR from L2/434 chromosomal DNA using primers #1212 and #1222, and #1221 and #1216, respectively. The DNA products were then fused by overlapping PCR using primers #1212 and #1216. The final DNA product was digested with NdeI-BamHI and ligated into those sites of pGBKT7.	4	This Work
pFA159	Y2H vector expressing Gal4BD-myc-$CT249_{\Delta TMD}$. Nucleotides 1 to 150, and 292 to 1053 of *ct249* were amplified by PCR from L2/434 chromosomal DNA using primers #943 and #1230, and #1229 and	4	This Work

Plasmid	Description/ Construction	Chapter	Reference
	#952, respectively. The DNA products were then fused by overlapping PCR using primers #943 and #952. The final DNA product was digested with NdeI-BamHI and ligated into those sites of pGBKT7.		
pCT288Int18	pGADT7 derivative plasmid expressing Gal4AD-HA-CCDC146$_{692-955}$ that was recovered from the Y2H screen using a CT288 fragment as bait.	4	This Work
pCT228Int3	pGADT7 derivative plasmid expressing Gal4AD-HA-LCA5L$_{5-279}$ that was recovered from the Y2H screen using a CT228 fragment as bait.	4	This Work
pCT249Int4	pGADT7 derivative plasmid expressing Gal4AD-HA-USP11$_{678-973}$ that was recovered from the Y2H screen using a CT249 fragment as bait.	4	This Work
pFA178	Y2H vector expressing Gal4BD-myc-CT288$_{89-241}$. Nucleotides 265 to 723 of *ct288* were amplified by PCR from L2/434 chromosomal DNA using primers #944 and #1474. The DNA product was digested with NdeI-BamHI and ligated into those sites of pGBKT7.	4	This Work
pFA179	Y2H vector expressing Gal4BD-myc-CT288$_{292-563}$. Nucleotides 874 to 1689 of *ct288* were amplified by PCR from L2/434 chromosomal DNA using primers #1475 and #954. The DNA product was digested with NdeI-BamHI and ligated into those sites of pGBKT7.	4	This Work
pFA184	Y2H vector expressing Gal4BD-myc-CT288$_{\Delta N \Delta TMD}$. Nucleotides 265 to 723, and 874 to 1689 of *ct288* were amplified by PCR from C/TW3 chromosomal DNA using primers #944 and #863, and #862 and #954, respectively. The DNA products were then fused by overlapping PCR using primers #944 and #954. The final DNA product was digested with NdeI-BamHI and ligated into those sites of pGBKT7.	4	This Work
pFA185	Y2H vector expressing Gal4BD-myc-CT288$_{\Delta N \Delta TMD}$. Nucleotides 265 to 723, and 874 to 1689 of *ct288* were amplified by PCR from E/Bour chromosomal DNA	4	This Work

Plasmid	Description/ Construction	Chapter	Reference
	using primers #944 and #863, and #862 and #954, respectively. The DNA products were then fused by overlapping PCR using primers #944 and #954. The final DNA product was digested with NdeI-BamHI and ligated into those sites of pGBKT7		
pFA139	pEGFP-C1 derivative plasmid encoding EGFP-CT288$_{\Delta N \Delta TMD}$. Nucleotides 265 to 723, and 874 to 1689 of *ct288* were amplified by PCR from L2/434 chromosomal DNA using primers #944 and #863, and #862 and #954, respectively. The DNA products were then fused by overlapping PCR using primers #944 and #954. The final DNA product was digested with KpnI-BamHI and ligated into those sites of pEGFP-C1.	4	This Work
pFA197	pEF6/Myc-His C derivative plasmid encoding CT288$_{\Delta N \Delta TMD}$-HA. Nucleotides 265 to 723, and 874 to 1689 of *ct288* were amplified by PCR from pSVP255 DNA using primers #1733 and #863, and #862 and #1734, respectively. The DNA products were then fused by overlapping PCR using primers #1733 and #1734. The final DNA product was digested with BamHI-EcoRI and ligated into those sites of pEF6/Myc-His C.	4	This Work
pFA164	pEGFP-C1 derivative plasmid encoding EGFP-CCDC146$_{FL}$. Nucleotides 1 to 2865 of *CCDC146* were amplified by PCR from pCCDC146 DNA using primers #1350 and #1351. The DNA product was digested with XhoI-EcoRI and ligated into those sites of pEGFP-C1.	4	This Work
pFA196	pEGFP-C1 derivative plasmid encoding EGFP-CCDC146$_{692\text{-}955}$. Nucleotides 2073 to 2865 of *CCDC146* were amplified by PCR from pCCDC146 DNA using primers #1730 and #1351. The DNA product was digested with XhoI-EcoRI and ligated into those sites of pEGFP-C1.	4	This Work
pFA167	pEF6/Myc-His C derivative plasmid encoding CCDC146$_{FL}$-HA. Nucleotides 1 to 2865 of *CCDC146* were amplified by PCR from pCCDC146 DNA using primers	4	This Work

Plasmid	Description/ Construction	Chapter	Reference
	#1352 and #1355. The DNA product was digested with EcoRI-NotI and ligated into those sites of pEF6/Myc-His C.		
pFA168	pEF6/Myc-His C derivative plasmid encoding $CCDC146_{692-955}$-HA. Nucleotides 2073 to 2865 of *CCDC146* were amplified by PCR from pCCDC146 DNA using primers #1354 and #1355. The DNA product was digested with EcoRI-NotI and ligated into those sites of pEF6/Myc-His C.	4	This Work

Agaisse, H., and Derré, I. (2013) A *C. trachomatis* Cloning Vector and the Generation of *C. trachomatis* Strains Expressing Fluorescent Proteins under the Control of a *C. trachomatis* Promoter. *PLoS One* **8**.

Charpentier, X., and Oswald, E. (2004) Identification of the secretion and translocation domain of the enteropathogenic and enterohemorrhagic *Escherichia coli* effector Cif, using TEM-1 beta-lactamase as a new fluorescence-based reporter. *J Bacteriol* **186**: 5486–5495.

Marenne, M.N., Journet, L., Mota, L.J., and Cornelis, G.R. (2003) Genetic analysis of the formation of the Ysc-Yop translocation pore in macrophages by *Yersinia enterocolitica*: Role of LcrV, YscF and YopN. *Microb Pathog* **35**: 243–258.

Sorg, I., Wagner, S., Amstutz, M., Müller, S. a, Broz, P., Lussi, Y., et al. (2007) YscU recognizes translocators as export substrates of the *Yersinia* injectisome. *EMBO J* **26**: 3015–3024.

Table A.3 – Accession numbers of *C. trachomatis* genomic and *inc* gene sequences used in this work.

Strain	Genome and *inc* accession numbers	Reference
A/Har13	CP000051	(Carlson et al., 2005)
A/2497	FM872306	(Seth-Smith et al., 2009)
A/363	HE601796	(Harris et al., 2012)
A/5291	HE601810	(Harris et al., 2012)
A/7249	HE601797	(Harris et al., 2012)
B/TZ1A4828/OT	FM872307	(Seth-Smith et al., 2009)
B/Jali20	FM872308	(Seth-Smith et al., 2009)
D/UW3/CX	AE001273	(Stephens et al., 1998)
D/SotonD1	HE601798	(Harris et al., 2012)
D/SotonD5	HE601799	(Harris et al., 2012)
D/SotonD6	HE601800	(Harris et al., 2012)
E/Bour	HE601870	(Harris et al., 2012)
E/SW2	FN652779	(Unemo et al., 2010)
E/SW3	HE601801	(Harris et al., 2012)
E/SotonE4	HE601802	(Harris et al., 2012)
E/SotonE8	HE601803	(Harris et al., 2012)
E/11023	CP001890	(Jeffrey et al., 2010)
E/150	CP001886	(Jeffrey et al., 2010)
F/SW4	HE601804	(Harris et al., 2012)
F/SW5	HE601805	(Harris et al., 2012)
F/Soton3	HE601806	(Harris et al., 2012)
G/9301	CP001930	(Jeffrey et al., 2010)
G/9768	CP001887	(Jeffrey et al., 2010)
G/11074	CP001889	(Jeffrey et al., 2010)
G/11222	CP001888	(Jeffrey et al., 2010)
G/SotonG1	HE601807	(Harris et al., 2012)
Ia/Sotonia1	HE601808	(Harris et al., 2012)
Ia/Sotonia3	HE601809	(Harris et al., 2012)
J/6276	ABYD01000001	(Jeffrey et al., 2010)
K/SotonK1	HE601794	(Harris et al., 2012)
L1/440/LN	HE601950	(Harris et al., 2012)
L1/1322/p2	HE601951	(Harris et al., 2012)
L1/115	HE601952	(Harris et al., 2012)
L1/224	HE601953	(Harris et al., 2012)
L2/434	AM884176	(Thomson et al., 2008)
L2/25567R	HE601954	(Harris et al., 2012)
L2b/UCH-1	AM884177	(Thomson et al., 2008)
L2b/8200/07	HE601795	(Harris et al., 2012)
L2b/UCH-2	HE601956	(Harris et al., 2012)
L2b/Canada1	HE601963	(Harris et al., 2012)
L2b/Canada2	HE601957	(Harris et al., 2012)
L2b/LST	HE601958	(Harris et al., 2012)
L2b/CV204	HE601960	(Harris et al., 2012)

Strain	Genome and *inc* accession numbers	Reference
L2b/795	HE601949	(Harris et al., 2012)
L2b/Ams1	HE601959	(Harris et al., 2012)
L2b/Ams2	HE601961	(Harris et al., 2012)
L2b/Ams3	HE601962	(Harris et al., 2012)
L2b/Ams4	HE601964	(Harris et al., 2012)
L2b/Ams5	HE601965	(Harris et al., 2012)
L2c	CP002024	(Somboonna et al., 2011)
L3/404/LN	HE601955	(Harris et al., 2012)

Carlson, J.H., Porcella, S.F., Mcclarty, G., and Caldwell, H.D. (2005) Comparative Genomic Analysis of *Chlamydia trachomatis* Oculotropic and Genitotropic Strains. *Infect Immun* **73**: 6407–6418.

Harris, S.R., Clarke, I.N., Seth-Smith, H.M.B., Solomon, A.W., Cutcliffe, L.T., Marsh, P., et al. (2012) Whole-genome analysis of diverse Chlamydia trachomatis strains identifies phylogenetic relationships masked by current clinical typing. *Nat Genet* **44**: 413–419.

Jeffrey, B.M., Suchland, R.J., Quinn, K.L., Davidson, J.R., Stamm, W.E., and Rockey, D.D. (2010) Genome sequencing of recent clinical *Chlamydia trachomatis* strains identifies loci associated with tissue tropism and regions of apparent recombination. *Infect Immun* **78**: 2544–2553.

Seth-Smith, H.M.B., Harris, S.R., Persson, K., Marsh, P., Barron, A., Bignell, A., et al. (2009) Co-evolution of genomes and plasmids within *Chlamydia trachomatis* and the emergence in Sweden of a new variant strain. *BMC Genomics* **10**: 239.

Somboonna, N., Wan, R., Ojcius, D.M., Pettengill, M. a, Joseph, S.J., Chang, A., et al. (2011) Hypervirulent *Chlamydia trachomatis* Clinical Strain Is a Recombinant between Lymphogranuloma Venereum (L2) and D Lineages. *MBio* **2**: 1–12.

Stephens, R.S., Kalman, S., Lammel, C., Fan, J., Marathe, R., Aravind, L., et al. (1998) Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*. *Science* **282**: 754–759.

Thomson, N.R., Holden, M.T.G., Carder, C., Lennard, N., Lockey, S.J., Marsh, P., et al. (2008) *Chlamydia trachomatis*: genome sequence analysis of lymphogranuloma venereum isolates. *Genome Res* **18**: 161–171.

Unemo, M., Seth-Smith, H.M.B., Cutcliffe, L.T., Skilton, R.J., Barlow, D., Goulding, D., et al. (2010) The Swedish new variant of *Chlamydia trachomatis*: Genome sequence, morphology, cell tropism and phenotypic characterization. *Microbiology* **156**: 1394–1404.

Table A.4 – Molecular evolution analyses of *inc*, *pmp*, and housekeeping (HK) genes among all 51 *C. trachomatis* strains[a].

All strains (ocular, urogenital, and LGV)								
Gene	**d_N**	**SEM**	**d_S**	**SEM**	**d_N/d_S**	**SEM**	**Z-test of selection**	
							positive	**neutral**
inc genes								
ct005	0,0018	0,0009	0,0084	0,0034	0,2104	0,1346	ND	ND
ct006	0,0019	0,0014	0,0002	0,0002	9,6440	11,5124	0,1136	ND
ct036	0,0025	0,0011	0,0044	0,0021	0,5787	0,3661	ND	ND
ct058	0,0103	0,0022	0,0076	0,0031	1,3412	0,6151	0,2470	ND
ct101	0,0025	0,0011	0,0044	0,0021	0,5787	0,3665	ND	ND
ct115	0,0296	0,0079	0,0262	0,0097	1,1287	0,5170	0,3669	ND
ct116	0,0282	0,0068	0,0082	0,0042	3,4243	1,9470	0,0035	0,0063
ct117	0,0188	0,0060	0,0088	0,0055	2,1380	1,4889	0,0870	ND
ct118	0,0149	0,0042	0,0030	0,0029	5,0151	5,0390	0,0150	0,0290
ct119	0,0087	0,0024	0,0033	0,0020	2,6132	1,7436	0,0250	0,0480
ct134	0,0028	0,0017	0,0087	0,0054	0,3160	0,2723	ND	ND
ct135	0,0043	0,0015	0,0067	0,0032	0,6353	0,3720	ND	ND
ct147	0,0088	0,0011	0,0066	0,0013	1,3417	0,3109	0,0830	ND
ct164	0,0030	0,0028	0,0000	0,0000	ND	ND	0,1419	ND
ct179	0,0036	0,0019	0,0040	0,0029	0,9096	0,8229	ND	ND
ct192	0,0106	0,0028	0,0093	0,0038	1,1456	0,5606	0,3880	ND
ct195	0,0043	0,0013	0,0086	0,0031	0,5042	0,2346	ND	ND
ct196	0,0043	0,0029	0,0071	0,0053	0,6020	0,6127	ND	ND
ct214	0,0064	0,0015	0,0064	0,0028	1,0094	0,5038	0,4883	ND
ct222	0,0197	0,0075	0,0110	0,0058	1,7925	1,1701	0,0240	0,0450
ct223	0,0278	0,0047	0,0065	0,0045	4,2470	3,0401	0,0026	0,0050
ct224	0,0043	0,0024	0,0018	0,0016	2,4013	2,4406	0,1639	ND
ct225	0,0013	0,0010	0,0043	0,0039	0,3149	0,3740	ND	ND
ct226	0,0049	0,0023	0,0042	0,0030	1,1754	1,0150	0,4224	ND
ct227	0,0052	0,0028	0,0017	0,0017	3,1477	3,5919	0,1405	ND
ct228	0,0108	0,0034	0,0023	0,0024	4,6703	4,9840	0,0210	0,0480
ct229	0,0263	0,0047	0,0055	0,0036	4,7722	3,2168	0,0002	0,0010
ct232	0,0088	0,0036	0,0064	0,0048	1,3802	1,1794	0,3201	ND
ct233	0,0069	0,0024	0,0125	0,0057	0,5492	0,3165	ND	ND
ct249	0,0211	0,0078	0,0000	0,0000	ND	ND	0,0038	0,0050
ct288	0,0117	0,0021	0,0045	0,0015	2,6166	0,9920	0,0022	0,0040

All strains (ocular, urogenital, and LGV)								
Gene	**d_N**	**SEM**	**d_S**	**SEM**	**d_N/d_S**	**SEM**	**Z-test of selection**	
							positive	**neutral**
ct300	0,0026	0,0020	0,0098	0,0062	0,2682	0,2621	ND	ND
ct345	0,0037	0,0023	0,0023	0,0024	1,6120	1,9630	0,3344	ND
ct357	0,0047	0,0029	0,0057	0,0044	0,8135	0,8090	ND	ND
ct358	0,0232	0,0070	0,0241	0,0082	0,9597	0,4369	ND	ND
ct365	0,0054	0,0014	0,0042	0,0016	1,2691	0,5935	0,2977	ND
ct383	0,0053	0,0020	0,0011	0,0010	4,8771	4,7082	0,0320	0,0630
ct440	0,0021	0,0020	0,0000	0,0000	ND	ND	0,1680	ND
ct442	0,0231	0,0053	0,0140	0,0079	1,6499	1,0024	0,0741	ND
ct449	0,0040	0,0028	0,0000	0,0000	ND	ND	0,0784	ND
ct483	0,0003	0,0003	0,0072	0,0048	0,0431	0,0516	ND	ND
ct484	0,0008	0,0006	0,0074	0,0039	0,1083	0,0988	ND	ND
ct565	0,0019	0,0018	0,0058	0,0035	0,3341	0,3630	ND	ND
ct618	0,0148	0,0033	0,0314	0,0085	0,4707	0,1649	ND	ND
ct728	0,0021	0,0012	0,0045	0,0025	0,4619	0,3808	ND	ND
ct789	0,0000	0,0000	0,0057	0,0064	0,0000	ND	ND	ND
ct813	0,0164	0,0036	0,0040	0,0024	4,1404	2,6649	0,0028	0,0063
ct850	0,0019	0,0009	0,0035	0,0017	0,5517	0,3792	ND	ND
pmp genes								
pmpA	0,0008	0,0003	0,0018	0,0008	0,4504	0,2703	ND	ND
pmpB	0,0083	0,0009	0,0051	0,0013	1,6211	0,4442	0,0220	0,0431
pmpC	0,0058	0,0008	0,0040	0,0009	1,4323	0,3684	0,0653	ND
pmpD	0,0024	0,0006	0,0039	0,0010	0,6290	0,2118	ND	ND
pmpE	0,0148	0,0017	0,0484	0,0052	0,3066	0,0486	ND	ND
pmpF	0,0415	0,0031	0,1595	0,0093	0,2603	0,0248	ND	ND
pmpG	0,0071	0,0014	0,0174	0,0029	0,4049	0,1040	ND	ND
pmpH	0,0166	0,0019	0,0923	0,0079	0,1794	0,0260	ND	ND
pmpI	0,0031	0,0008	0,0094	0,0020	0,3324	0,1100	ND	ND
HK genes								
yraL (ct048)	0,0057	0,0018	0,0078	0,0033	0,7375	0,3916	ND	ND
fer (ct059)	0,0048	0,0034	0,0242	0,0112	0,1992	0,1677	ND	ND
accD (ct293)	0,0035	0,0015	0,0433	0,0096	0,0807	0,0392	ND	ND
karG (ct675)	0,0063	0,0017	0,0560	0,0092	0,1124	0,0359	ND	ND
rrf (ct677)	0,0057	0,0025	0,0930	0,0163	0,0613	0,0288	ND	ND
pyrH (ct678)	0,0039	0,0016	0,0339	0,0074	0,1152	0,0536	ND	ND

All strains (ocular, urogenital, and LGV)								
Gene	**d_N**	**SEM**	**d_S**	**SEM**	**d_N/d_S**	**SEM**	**Z-test of selection**	
							positive	**neutral**
tsf (ct679)	0,0099	0,0024	0,1036	0,0140	0,0955	0,0263	ND	ND
rpsB (ct680)	0,0168	0,0036	0,1051	0,0131	0,1595	0,0398	ND	ND
yfO_1 (ct687)	0,0015	0,0009	0,0226	0,0054	0,0673	0,0417	ND	ND

[a]Non-synonymous (d_N) and synonymous (d_S) substitutions between *inc*, *pmp*, and housekeeping (HK) genes among all 51 *C. trachomatis* strains, and test of selection using the codon-based Z test using MEGA5 (Tamura et al., 2011) as described in Materials and Methods. When dN/dS > 1, we tested for positive selection. If the results of the codon based Z-tests of selection were significant ($p < 0.05$), we further tested for lack of neutrality ($d_N \neq d_S$). Genes were considered likely under putative positive selection, only if both positive selection and neutrality tests were significant. The nucleotide sequences were obtained from the genomes of the 51 *C. trachomatis* strains listed in Table A.3.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**: 2731–2739.

Table A.5 – Molecular evolution analyses of *inc*, *pmp*, and housekeeping (HK) genes among all 51 *C. trachomatis* strains minus urogenital strains[a].

All strains minus urogenital (ocular and LGV)								
Gene	**d_N**	**SEM**	**d_S**	**SEM**	**d_N/d_S**	**SEM**	**Z-test of selection**	
							positive	**neutral**
inc genes								
ct005	ND	ND	ND	ND	ND	ND	ND	ND
ct006	ND	ND	ND	ND	ND	ND	ND	ND
ct036	ND	ND	ND	ND	ND	ND	ND	ND
ct058	ND	ND	ND	ND	ND	ND	ND	ND
ct101	ND	ND	ND	ND	ND	ND	ND	ND
ct115	ND	ND	ND	ND	ND	ND	ND	ND
ct116	0,0168	0,0048	0,0063	0,0039	2,6526	1,7899	0,0324	0,0650
ct117	ND	ND	ND	ND	ND	ND	ND	ND
ct118	0,0118	0,0038	0,0026	0,0024	4,4573	4,3758	0,0166	0,0361
ct119	0,0080	0,0023	0,0018	0,0016	4,5316	4,4106	0,0024	0,0045
ct134	ND	ND	ND	ND	ND	ND	ND	ND
ct135	ND	ND	ND	ND	ND	ND	ND	ND
ct147	ND	ND	ND	ND	ND	ND	ND	ND
ct164	ND	ND	ND	ND	ND	ND	ND	ND
ct179	ND	ND	ND	ND	ND	ND	ND	ND
ct192	ND	ND	ND	ND	ND	ND	ND	ND
ct195	ND	ND	ND	ND	ND	ND	ND	ND
ct196	ND	ND	ND	ND	ND	ND	ND	ND
ct214	ND	ND	ND	ND	ND	ND	ND	ND
ct222	0,0160	0,0060	0,0101	0,0054	1,5857	1,0330	0,0809	ND
ct223	0,0229	0,0042	0,0046	0,0034	4,9942	3,8494	0,0038	0,0092
ct224	ND	ND	ND	ND	ND	ND	ND	ND
ct225	ND	ND	ND	ND	ND	ND	ND	ND
ct226	ND	ND	ND	ND	ND	ND	ND	ND
ct227	ND	ND	ND	ND	ND	ND	ND	ND
ct228	0,0075	0,0025	0,0015	0,0016	4,8536	5,1520	0,0210	0,0410
ct229	0,0174	0,0036	0,0037	0,0024	4,7103	3,2301	0,0013	0,0035
ct232	ND	ND	ND	ND	ND	ND	ND	ND
ct233	ND	ND	ND	ND	ND	ND	ND	ND
ct249	0,0159	0,0056	0,0000	0,0000	ND	ND	0,0044	0,0062
ct288	0,0069	0,0014	0,0026	0,0011	2,6315	1,2039	0,0123	0,0283

All strains minus urogenital (ocular and LGV)								
Gene	**d_N**	**SEM**	**d_S**	**SEM**	**d_N/d_S**	**SEM**	**Z-test of selection**	
							positive	**neutral**
ct300	ND	ND	ND	ND	ND	ND	ND	ND
ct345	ND	ND	ND	ND	ND	ND	ND	ND
ct357	ND	ND	ND	ND	ND	ND	ND	ND
ct358	ND	ND	ND	ND	ND	ND	ND	ND
ct365	ND	ND	ND	ND	ND	ND	ND	ND
ct383	ND	ND	ND	ND	ND	ND	ND	ND
ct440	ND	ND	ND	ND	ND	ND	ND	ND
ct442	ND	ND	ND	ND	ND	ND	ND	ND
ct449	ND	ND	ND	ND	ND	ND	ND	ND
ct483	ND	ND	ND	ND	ND	ND	ND	ND
ct484	ND	ND	ND	ND	ND	ND	ND	ND
ct565	ND	ND	ND	ND	ND	ND	ND	ND
ct618	ND	ND	ND	ND	ND	ND	ND	ND
ct728	ND	ND	ND	ND	ND	ND	ND	ND
ct789	ND	ND	ND	ND	ND	ND	ND	ND
ct813	0,0114	0,0028	0,0020	0,0015	5,5693	4,2519	0,0020	0,0032
ct850	ND	ND	ND	ND	ND	ND	ND	ND
pmp genes								
pmpA	ND	ND	ND	ND	ND	ND	ND	ND
pmpB	ND	ND	ND	ND	ND	ND	ND	ND
pmpC	ND	ND	ND	ND	ND	ND	ND	ND
pmpD	ND	ND	ND	ND	ND	ND	ND	ND
pmpE	ND	ND	ND	ND	ND	ND	ND	ND
pmpF	ND	ND	ND	ND	ND	ND	ND	ND
pmpG	ND	ND	ND	ND	ND	ND	ND	ND
pmpH	ND	ND	ND	ND	ND	ND	ND	ND
pmpI	ND	ND	ND	ND	ND	ND	ND	ND
HK genes								
yraL (ct048)	ND	ND	ND	ND	ND	ND	ND	ND
fer (ct059)	ND	ND	ND	ND	ND	ND	ND	ND
accD (ct293)	ND	ND	ND	ND	ND	ND	ND	ND
karG (ct675)	ND	ND	ND	ND	ND	ND	ND	ND
rrf (ct677)	ND	ND	ND	ND	ND	ND	ND	ND
pyrH (ct678)	ND	ND	ND	ND	ND	ND	ND	ND

All strains minus urogenital (ocular and LGV)								
Gene	**d_N**	**SEM**	**d_S**	**SEM**	**d_N/d_S**	**SEM**	**Z-test of selection**	
							positive	**neutral**
tsf (ct679)	ND	ND	ND	ND	ND	ND	ND	ND
rpsB (ct680)	ND	ND	ND	ND	ND	ND	ND	ND
yfO_1 (ct687)	ND	ND	ND	ND	ND	ND	ND	ND

[a]Non-synonymous (d_N) and synonymous (d_S) substitutions between *inc*, *pmp*, and housekeeping (HK) genes among all *C. trachomatis* strains minus urogenital strains, and test of selection using the codon-based Z test using MEGA5 (Tamura et al., 2011) as described in Materials and Methods. When dN/dS > 1, we tested for positive selection. If the results of the codon based Z-tests of selection were significant ($p < 0.05$), we further tested for lack of neutrality ($d_N \neq d_S$). Genes were considered likely under putative positive selection, only if both positive selection and neutrality tests were significant. The nucleotide sequences were obtained from the genomes of the 51 *C. trachomatis* strains listed in Table A.3.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**: 2731–2739.

Table A.6 – Molecular evolution analyses of *inc*, *pmp*, and housekeeping (HK) genes among all 51 *C. trachomatis* strains minus ocular strains[a].

All strains minus ocular (urogenital and LGV)								
Gene	**d_N**	**SEM**	**d_S**	**SEM**	**d_N/d_S**	**SEM**	**Z-test of selection**	
							positive	**Neutral**
inc genes								
ct005	ND	ND	ND	ND	ND	ND	ND	ND
ct006	ND	ND	ND	ND	ND	ND	ND	ND
ct036	ND	ND	ND	ND	ND	ND	ND	ND
ct058	ND	ND	ND	ND	ND	ND	ND	ND
ct101	ND	ND	ND	ND	ND	ND	ND	ND
ct115	ND	ND	ND	ND	ND	ND	ND	ND
ct116	0,0280	0,0068	0,0068	0,0042	4,1397	2,7567	0,0015	0,0035
ct117	ND	ND	ND	ND	ND	ND	ND	ND
ct118	0,0157	0,0045	0,0031	0,0031	5,0004	5,1294	0,0120	0,0286
ct119	0,0083	0,0025	0,0037	0,0022	2,2331	1,4850	0,0560	ND
ct134	ND	ND	ND	ND	ND	ND	ND	ND
ct135	ND	ND	ND	ND	ND	ND	ND	ND
ct147	ND	ND	ND	ND	ND	ND	ND	ND
ct164	ND	ND	ND	ND	ND	ND	ND	ND
ct179	ND	ND	ND	ND	ND	ND	ND	ND
ct192	ND	ND	ND	ND	ND	ND	ND	ND
ct195	ND	ND	ND	ND	ND	ND	ND	ND
ct196	ND	ND	ND	ND	ND	ND	ND	ND
ct214	ND	ND	ND	ND	ND	ND	ND	ND
ct222	0,0203	0,0077	0,0096	0,0057	2,1117	1,4924	0,0052	0,0133
ct223	0,0254	0,0048	0,0060	0,0043	4,1962	3,0885	0,0033	0,0065
ct224	ND	ND	ND	ND	ND	ND	ND	ND
ct225	ND	ND	ND	ND	ND	ND	ND	ND
ct226	ND	ND	ND	ND	ND	ND	ND	ND
ct227	ND	ND	ND	ND	ND	ND	ND	ND
ct228	0,0106	0,0035	0,0024	0,0024	4,4774	4,7061	0,0300	0,0560
ct229	0,0262	0,0052	0,0050	0,0034	5,2698	3,6992	0,0023	0,0083
ct232	ND	ND	ND	ND	ND	ND	ND	ND
ct233	ND	ND	ND	ND	ND	ND	ND	ND
ct249	0,0205	0,0077	0,0000	0,0000	ND	ND	0,0048	0,0077
ct288	0,0118	0,0021	0,0044	0,0015	2,7030	1,0673	0,0026	0,0057

All strains minus ocular (urogenital and LGV)								
Gene	**d_N**	**SEM**	**d_S**	**SEM**	**d_N/d_S**	**SEM**	**Z-test of selection**	
							positive	**Neutral**
ct300	ND	ND	ND	ND	ND	ND	ND	ND
ct345	ND	ND	ND	ND	ND	ND	ND	ND
ct357	ND	ND	ND	ND	ND	ND	ND	ND
ct358	ND	ND	ND	ND	ND	ND	ND	ND
ct365	ND	ND	ND	ND	ND	ND	ND	ND
ct383	ND	ND	ND	ND	ND	ND	ND	ND
ct440	ND	ND	ND	ND	ND	ND	ND	ND
ct442	ND	ND	ND	ND	ND	ND	ND	ND
ct449	ND	ND	ND	ND	ND	ND	ND	ND
ct483	ND	ND	ND	ND	ND	ND	ND	ND
ct484	ND	ND	ND	ND	ND	ND	ND	ND
ct565	ND	ND	ND	ND	ND	ND	ND	ND
ct618	ND	ND	ND	ND	ND	ND	ND	ND
ct728	ND	ND	ND	ND	ND	ND	ND	ND
ct789	ND	ND	ND	ND	ND	ND	ND	ND
ct813	0,0168	0,0037	0,0042	0,0025	3,9847	2,5200	0,0027	0,0084
ct850	ND	ND	ND	ND	ND	ND	ND	ND
pmp genes								
pmpA	ND	ND	ND	ND	ND	ND	ND	ND
pmpB	ND	ND	ND	ND	ND	ND	ND	ND
pmpC	ND	ND	ND	ND	ND	ND	ND	ND
pmpD	ND	ND	ND	ND	ND	ND	ND	ND
pmpE	ND	ND	ND	ND	ND	ND	ND	ND
pmpF	ND	ND	ND	ND	ND	ND	ND	ND
pmpG	ND	ND	ND	ND	ND	ND	ND	ND
pmpH	ND	ND	ND	ND	ND	ND	ND	ND
pmpI	ND	ND	ND	ND	ND	ND	ND	ND
HK genes								
yraL (ct048)	ND	ND	ND	ND	ND	ND	ND	ND
fer (ct059)	ND	ND	ND	ND	ND	ND	ND	ND
accD (ct293)	ND	ND	ND	ND	ND	ND	ND	ND
karG (ct675)	ND	ND	ND	ND	ND	ND	ND	ND
rrf (ct677)	ND	ND	ND	ND	ND	ND	ND	ND
pyrH (ct678)	ND	ND	ND	ND	ND	ND	ND	ND

All strains minus ocular (urogenital and LGV)								
Gene	**d_N**	**SEM**	**d_S**	**SEM**	**d_N/d_S**	**SEM**	**Z-test of selection**	
							positive	**Neutral**
tsf (ct679)	ND	ND	ND	ND	ND	ND	ND	ND
rpsB (ct680)	ND	ND	ND	ND	ND	ND	ND	ND
yfO_1 (ct687)	ND	ND	ND	ND	ND	ND	ND	ND

[a]Non-synonymous (d_N) and synonymous (d_S) substitutions between *inc*, *pmp*, and housekeeping (HK) genes among all *C. trachomatis* strains minus ocular strains, and test of selection using the codon-based Z test using MEGA5 (Tamura et al., 2011) as described in Materials and Methods. When dN/dS > 1, we tested for positive selection. If the results of the codon based Z-tests of selection were significant ($p < 0.05$), we further tested for lack of neutrality ($d_N \neq d_S$). Genes were considered likely under putative positive selection, only if both positive selection and neutrality tests were significant. The nucleotide sequences were obtained from the genomes of the 51 *C. trachomatis* strains listed in Table A.3.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**: 2731–2739.

Table A.7 – Molecular evolution analyses of *inc*, *pmp*, and housekeeping (HK) genes among all 51 *C. trachomatis* strains minus LGV strains[a].

All strains minus LGV (ocular and urogenital)								
Gene	**d_N**	**SEM**	**d_S**	**SEM**	**d_N/d_S**	**SEM**	**Z-test of selection**	
							positive	**neutral**
inc genes								
ct005	ND	ND	ND	ND	ND	ND	ND	ND
ct006	ND	ND	ND	ND	ND	ND	ND	ND
ct036	ND	ND	ND	ND	ND	ND	ND	ND
ct058	ND	ND	ND	ND	ND	ND	ND	ND
ct101	ND	ND	ND	ND	ND	ND	ND	ND
ct115	ND	ND	ND	ND	ND	ND	ND	ND
ct116	0,0144	0,0045	0,0067	0,0040	2,1510	1,4507	0,1031	ND
ct117	ND	ND	ND	ND	ND	ND	ND	ND
ct118	0,0034	0,0020	0,0000	0,0000	ND	ND	0,0689	ND
ct119	0,0031	0,0013	0,0021	0,0016	1,4726	1,2506	0,3118	ND
ct134	ND	ND	ND	ND	ND	ND	ND	ND
ct135	ND	ND	ND	ND	ND	ND	ND	ND
ct147	ND	ND	ND	ND	ND	ND	ND	ND
ct164	ND	ND	ND	ND	ND	ND	ND	ND
ct179	ND	ND	ND	ND	ND	ND	ND	ND
ct192	ND	ND	ND	ND	ND	ND	ND	ND
ct195	ND	ND	ND	ND	ND	ND	ND	ND
ct196	ND	ND	ND	ND	ND	ND	ND	ND
ct214	ND	ND	ND	ND	ND	ND	ND	ND
ct222	0,0008	0,0006	0,0025	0,0026	0,3078	0,3852	ND	ND
ct223	0,0110	0,0026	0,0004	0,0003	28,016	24,1101	0,0401	0,0952
ct224	ND	ND	ND	ND	ND	ND	ND	ND
ct225	ND	ND	ND	ND	ND	ND	ND	ND
ct226	ND	ND	ND	ND	ND	ND	ND	ND
ct227	ND	ND	ND	ND	ND	ND	ND	ND
ct228	0,0015	0,0009	0,0000	0,0000	ND	ND	0,0580	ND
ct229	0,0078	0,0024	0,0006	0,0005	12,070	10,8707	0,0679	ND
ct232	ND	ND	ND	ND	ND	ND	ND	ND
ct233	ND	ND	ND	ND	ND	ND	ND	ND
ct249	0,0087	0,0036	0,0000	0,0000	ND	ND	0,0101	0,0172
ct288	0,0059	0,0015	0,0038	0,0016	1,5551	0,7563	0,1648	ND

All strains minus LGV (ocular and urogenital)								
Gene	**d_N**	**SEM**	**d_S**	**SEM**	**d_N/d_S**	**SEM**	**Z-test of selection**	
							positive	**neutral**
ct300	ND	ND	ND	ND	ND	ND	ND	ND
ct345	ND	ND	ND	ND	ND	ND	ND	ND
ct357	ND	ND	ND	ND	ND	ND	ND	ND
ct358	ND	ND	ND	ND	ND	ND	ND	ND
ct365	ND	ND	ND	ND	ND	ND	ND	ND
ct383	ND	ND	ND	ND	ND	ND	ND	ND
ct440	ND	ND	ND	ND	ND	ND	ND	ND
ct442	ND	ND	ND	ND	ND	ND	ND	ND
ct449	ND	ND	ND	ND	ND	ND	ND	ND
ct483	ND	ND	ND	ND	ND	ND	ND	ND
ct484	ND	ND	ND	ND	ND	ND	ND	ND
ct565	ND	ND	ND	ND	ND	ND	ND	ND
ct618	ND	ND	ND	ND	ND	ND	ND	ND
ct728	ND	ND	ND	ND	ND	ND	ND	ND
ct789	ND	ND	ND	ND	ND	ND	ND	ND
ct813	0,0046	0,0018	0,0024	0,0020	1,8906	1,7002	0,2028	ND
ct850	ND	ND	ND	ND	ND	ND	ND	ND
pmp genes								
pmpA	ND	ND	ND	ND	ND	ND	ND	ND
pmpB	ND	ND	ND	ND	ND	ND	ND	ND
pmpC	ND	ND	ND	ND	ND	ND	ND	ND
pmpD	ND	ND	ND	ND	ND	ND	ND	ND
pmpE	ND	ND	ND	ND	ND	ND	ND	ND
pmpF	ND	ND	ND	ND	ND	ND	ND	ND
pmpG	ND	ND	ND	ND	ND	ND	ND	ND
pmpH	ND	ND	ND	ND	ND	ND	ND	ND
pmpI	ND	ND	ND	ND	ND	ND	ND	ND
HK genes								
yraL (ct048)	ND	ND	ND	ND	ND	ND	ND	ND
fer (ct059)	ND	ND	ND	ND	ND	ND	ND	ND
accD (ct293)	ND	ND	ND	ND	ND	ND	ND	ND
karG (ct675)	ND	ND	ND	ND	ND	ND	ND	ND
rrf (ct677)	ND	ND	ND	ND	ND	ND	ND	ND
pyrH (ct678)	ND	ND	ND	ND	ND	ND	ND	ND

All strains minus LGV (ocular and urogenital)								
Gene	**d_N**	**SEM**	**d_S**	**SEM**	**d_N/d_S**	**SEM**	**Z-test of selection**	
							positive	**neutral**
tsf (ct679)	ND	ND	ND	ND	ND	ND	ND	ND
rpsB (ct680)	ND	ND	ND	ND	ND	ND	ND	ND
yfO_1 (ct687)	ND	ND	ND	ND	ND	ND	ND	ND

[a]Non-synonymous (d_N) and synonymous (d_S) substitutions between *inc*, *pmp*, and housekeeping (HK) genes among all *C. trachomatis* strains minus LGV strains, and test of selection using the codon-based Z test using MEGA5 (Tamura et al., 2011) as described in Materials and Methods. When dN/dS > 1, we tested for positive selection. If the results of the codon based Z-tests of selection were significant ($p < 0.05$), we further tested for lack of neutrality ($d_N \neq d_S$). Genes were considered likely under putative positive selection, only if both positive selection and neutrality tests were significant. The nucleotide sequences were obtained from the genomes of the 51 *C. trachomatis* strains listed in Table A.3.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**: 2731–2739.

Table A.8 – mRNA levels of *inc* genes during the developmental cycle of *Chlamydia trachomatis* C/TW3 strain[a].

***Chlamydia trachomatis* C/TW3**														
	Mean and SEM values at different time points (h) post-infection (p.i.)												**Profile of expression**	**Peak of expression (time p.i)**
inc	**T2**		**T6**		**T12**		**T20**		**T30**		**T42**			
	Mean	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**		
ct005	24,66	11,12	16,45	5,34	13,83	5,59	7,65	0,22	24,63	10,95	70,20	13,26	Early-Late	T42
ct006	91,67	7,05	61,06	4,43	30,59	3,82	20,07	1,60	9,76	0,71	7,34	1,02	Early	T2
ct036	50,53	8,78	34,57	1,67	24,54	3,04	24,52	2,55	16,58	3,04	14,82	2,12	Early	T2
ct058	0,60	0,41	0,28	0,05	0,43	0,32	0,14	0,08	0,11	0,07	0,10	0,06	Early	T2
ct101	13,19	ND	18,46	4,59	41,32	8,36	52,79	4,28	30,16	ND	8,89	1,10	Mid	T20
ct115	427,48	184,83	250,63	92,39	163,71	23,87	81,98	19,32	31,18	14,73	17,10	3,04	Early	T2
ct116	575,31	65,26	435,69	78,80	403,62	116,39	264,47	77,90	100,28	40,59	48,84	1,50	Early	T2
ct117	331,73	11,70	249,55	28,31	368,20	18,41	222,35	51,12	69,31	26,01	34,48	ND	Early	T12
ct118	200,24	23,06	196,98	3,93	324,39	31,76	184,41	21,14	62,80	18,71	25,96	3,43	Early	T12
ct119	17,18	4,89	11,39	0,88	24,31	1,91	28,02	2,06	27,21	1,18	12,62	1,55	Mid	T20
ct134	151,81	20,91	145,51	31,48	133,52	82,07	27,41	11,79	12,67	3,69	11,64	2,81	Early	T2
ct135	156,85	24,41	116,14	31,36	94,77	44,70	24,40	7,41	16,81	6,66	11,26	2,41	Early	T2
ct147	200,31	15,13	123,79	14,20	76,45	9,23	16,03	0,59	13,07	0,83	19,67	0,64	Early	T2
ct164	6,26	ND	4,56	0,87	13,53	ND	5,21	0,98	5,76	0,36	5,08	0,38	Early	T12
ct179	22,78	ND	23,69	2,69	50,50	9,72	26,53	5,45	14,70	ND	13,10	1,66	Mid	T12
ct192	8,26	2,02	25,27	1,03	19,71	6,87	6,55	2,78	3,90	0,92	3,70	0,15	Early	T6
ct195	15,97	ND	13,24	2,08	16,55	0,67	12,82	2,13	14,63	1,86	6,91	0,93	Early	T12
ct196	41,82	10,23	31,63	16,93	19,98	10,20	7,54	2,07	6,13	2,38	5,17	0,18	Early	T2
ct214	21,80	4,55	19,07	5,89	23,80	12,77	12,08	2,42	42,08	0,58	44,04	13,33	Late	T42
ct222	7,48	4,30	3,64	1,72	7,26	1,51	11,74	5,37	7,56	2,19	5,80	0,18	Mid	T20
ct223	5,53	3,81	3,64	1,53	15,16	1,01	52,81	23,09	22,16	8,22	10,78	2,15	Mid	T20
ct224	12,05	ND	5,78	0,42	5,19	0,46	6,13	0,49	10,78	2,19	5,12	0,99	Early	T2
ct225	10,25	0,45	5,47	0,98	10,16	1,60	12,62	1,70	17,34	1,74	7,51	0,94	Early	T30
ct226	6,74	2,43	10,11	ND	19,41	1,14	14,54	2,44	12,31	1,46	6,96	0,96	Mid	T12
ct227	17,00	5,78	12,24	0,76	15,88	7,48	6,44	0,31	7,73	2,08	5,55	0,32	Early	T2
ct228	75,13	3,97	76,00	6,79	73,19	29,83	28,44	13,41	11,23	3,65	8,94	0,27	Early	T6
ct229	372,22	54,35	239,91	76,64	82,41	29,58	23,63	7,19	24,23	1,88	53,43	14,36	Early	T2
ct232	79,65	14,68	95,63	18,56	78,47	30,62	60,31	33,40	24,45	9,31	11,15	3,16	Early	T6
ct233	20,33	6,72	20,24	3,30	20,48	8,46	14,99	7,68	19,10	8,82	12,52	4,55	Early	T12

Chlamydia trachomatis C/TW3														
	Mean and SEM values at different time points (h) post-infection (p.i.)												**Profile of expression**	**Peak of expression (time p.i)**
inc	**T2**		**T6**		**T12**		**T20**		**T30**		**T42**			
	Mean	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**		
ct249	30,70	0,80	20,63	1,53	26,84	8,37	18,31	7,62	12,72	3,40	9,95	0,09	Early	T2
ct288	20,51	8,54	18,93	6,62	34,38	15,13	16,73	8,30	30,62	10,21	39,76	4,13	Early	T42
ct300	13,51	2,42	9,30	0,67	23,49	4,75	9,09	0,17	6,10	0,12	10,09	0,77	Early	T12
ct345	10,84	ND	8,64	2,54	23,25	ND	7,10	0,13	7,52	0,65	5,92	0,21	Mid	T12
ct357	75,87	3,78	64,74	2,02	45,12	6,94	21,44	0,88	20,07	0,27	30,73	2,24	Early	T2
ct358	60,13	6,58	58,77	2,01	107,26	13,67	40,50	3,47	19,95	1,04	26,56	0,60	Early	T12
ct365	8,05	0,58	4,36	0,43	8,18	4,54	5,80	0,92	15,69	2,01	24,31	4,05	Late	T42
ct383	39,53	12,33	37,70	9,15	47,55	6,08	38,29	5,79	15,93	7,16	11,41	1,53	Early	T12
ct440	11,01	4,43	10,55	0,86	13,77	3,60	9,16	3,54	6,91	1,11	5,37	0,07	Early	T12
ct442	7,69	2,87	7,35	3,90	5,26	2,32	2,92	0,01	7,52	2,65	10,07	1,41	Early	T42
ct449	31,43	6,97	27,75	3,52	31,75	11,12	8,75	2,25	5,07	0,33	4,52	0,28	Early	T12
ct483	15,40	1,68	11,44	1,72	20,66	4,38	17,68	0,39	13,93	2,35	9,10	0,70	Early	T12
ct484	9,31	2,60	6,58	0,59	18,28	1,32	11,34	0,73	8,47	0,31	4,57	0,25	Early	T12
ct565	92,29	9,79	45,42	5,62	39,20	1,43	26,97	3,64	199,71	83,87	304,06	58,56	Late	T42
ct618	275,96	6,34	275,62	21,44	451,48	28,23	264,33	9,19	50,71	1,42	33,89	0,45	Early	T12
ct728	17,62	2,63	40,01	2,37	119,14	9,26	91,03	7,68	68,44	2,64	33,07	1,08	Mid	T12
ct789	26,58	7,22	18,17	3,59	23,09	7,83	17,83	0,31	13,45	1,68	8,37	0,48	Early	T2
ct813	11,91	4,83	12,29	2,95	9,75	0,52	6,21	0,58	17,89	10,02	30,61	5,81	Late	T42
ct850	7,35	2,33	3,73	0,36	22,61	0,97	17,99	3,60	25,32	6,36	10,43	1,91	Mid	T30

[a]Real-time quantitative PCR (RT-qPCR) of mRNA levels of *inc* genes during the developmental cycle of *Chlamydia trachomatis* C/TW3 strain. The data shows for each *inc* at each time point the mean values ($x10^5$) ± SEM, normalized against the corresponding values of the *16SrRNA*, from at least two independent experiments (except where indicated not determined, ND).

Table A.9 –mRNA levels of *inc* genes during the developmental cycle of *Chlamydia trachomatis* E/Bour strain[a].

***Chlamydia trachomatis* E/Bour**														
	Mean and SEM values at different time points (h) post-infection (p.i.)												**Profile of expression**	**Peak of expression (time p.i)**
inc	**T2**		**T6**		**T12**		**T20**		**T30**		**T42**			
	Mean	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**		
ct005	77,40	38,80	44,51	20,08	38,61	15,55	18,75	13,07	118,42	3,70	123,01	24,30	Early-Late	T42
ct006	130,27	28,37	65,65	ND	45,01	1,19	23,85	2,49	10,76	2,22	25,65	2,17	Early	T2
ct036	92,07	32,06	67,23	ND	40,94	4,67	32,69	0,84	15,77	3,13	41,39	5,65	Early	T2
ct058	0,40	0,30	0,81	0,48	0,65	0,47	0,13	0,11	0,10	0,08	0,24	0,07	Early	T6
ct101	12,91	ND	22,71	2,61	40,44	1,60	35,47	0,66	10,37	2,09	23,75	2,08	Mid	T12
ct115	160,14	12,15	79,75	5,14	70,04	3,04	19,22	8,11	12,08	0,01	24,72	1,73	Early	T2
ct116	472,47	107,51	265,52	72,77	224,05	29,28	86,43	48,29	34,06	5,12	54,89	6,44	Early	T2
ct117	209,13	2,96	160,14	8,12	213,11	7,77	84,22	11,23	26,77	2,04	42,69	6,02	Early	T12
ct118	173,49	70,41	133,41	ND	177,53	31,36	92,39	2,44	24,57	6,20	47,30	15,77	Early	T12
ct119	4,48	1,14	12,49	ND	21,36	2,59	46,23	5,39	19,83	1,95	23,55	1,13	Mid	T20
ct134	297,41	130,60	173,09	79,97	89,70	35,53	26,25	19,96	16,74	5,23	48,29	17,54	Early	T2
ct135	272,79	116,21	134,17	51,39	88,81	23,86	25,79	16,41	17,86	4,81	39,89	6,52	Early	T2
ct147	138,02	6,12	73,36	10,37	82,72	5,10	10,82	0,25	33,42	2,18	51,36	3,35	Early	T2
ct164	5,92	2,42	7,94	1,62	8,26	1,50	9,46	0,46	5,30	0,29	18,02	2,55	Late	T42
ct179	18,84	7,41	34,25	ND	63,23	9,14	27,81	1,30	12,13	3,28	23,83	3,01	Mid	T12
ct192	4,26	1,69	7,10	3,83	4,35	0,83	2,91	1,12	4,32	0,02	18,30	3,03	Late	T42
ct195	13,64	3,92	16,56	0,83	18,21	0,83	21,90	2,70	13,52	1,05	20,23	3,57	Early	T20
ct196	65,34	28,14	33,40	20,10	15,81	6,70	8,77	6,24	6,39	1,42	23,78	1,66	Early	T2
ct214	16,41	6,80	11,79	4,02	10,54	2,61	12,21	4,89	48,09	0,35	55,17	8,76	Late	T42
ct222	4,86	3,07	6,91	5,37	7,39	3,12	8,82	3,62	6,87	0,97	20,01	1,78	Late	T42
ct223	14,51	12,39	8,03	5,53	22,92	1,94	52,41	21,45	22,44	6,27	29,87	0,03	Mid	T20
ct224	11,90	3,68	11,20	1,84	8,16	0,22	9,81	0,22	7,01	0,74	18,13	0,90	Early	T42
ct225	10,67	3,78	15,64	3,09	16,50	1,38	17,69	2,17	10,05	1,04	21,60	1,01	Early	T42
ct226	10,22	0,60	11,17	5,51	21,38	8,11	14,86	1,77	8,36	0,63	17,89	2,42	Mid	T12
ct227	28,37	13,85	36,93	20,77	25,19	9,74	8,83	5,74	7,37	0,69	26,62	2,14	Early	T6
ct228	210,25	59,54	187,71	66,67	199,68	24,15	53,40	37,36	18,50	0,49	36,11	8,07	Early	T2
ct229	491,19	117,14	174,29	66,33	47,15	6,64	12,96	5,21	59,03	4,57	125,67	49,47	Early	T2
ct232	145,79	83,39	70,00	25,56	44,49	10,94	21,31	14,14	11,90	4,85	24,62	2,97	Early	T2
ct233	22,09	11,00	20,98	11,39	15,75	5,83	18,52	9,21	21,38	11,68	33,33	5,39	Early	T42

***Chlamydia trachomatis* E/Bour**														
	Mean and SEM values at different time points (h) post-infection (p.i.)												**Profile of expression**	**Peak of expression (time p.i)**
inc	**T2**		**T6**		**T12**		**T20**		**T30**		**T42**			
	Mean	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**		
ct249	66,11	25,16	49,62	20,53	20,18	3,75	14,42	6,51	16,65	5,58	31,82	3,18	Early	T2
ct288	28,73	11,13	21,66	13,94	20,40	4,95	12,55	6,27	61,62	2,03	52,32	0,41	Late	T30
ct300	7,43	1,38	15,56	1,42	15,72	1,12	9,36	1,09	8,62	1,49	25,25	3,30	Early	T42
ct345	4,83	1,41	10,81	2,20	12,65	2,42	21,44	0,85	13,84	0,24	19,62	1,26	Mid	T20
ct357	63,83	9,81	36,62	6,07	32,74	4,06	23,37	1,27	21,56	4,03	29,74	4,07	Early	T2
ct358	46,05	10,51	40,75	2,87	50,84	2,64	19,54	0,68	19,21	0,35	36,33	1,44	Early-Late	T12
ct365	5,22	0,23	9,51	0,72	7,07	1,23	7,20	1,32	16,55	2,01	27,67	3,81	Late	T42
ct383	41,27	8,12	24,65	1,16	38,91	3,49	18,87	3,11	13,62	1,65	18,89	2,73	Early	T2
ct440	14,23	5,59	13,72	5,34	20,34	7,20	8,57	4,59	6,24	0,78	18,07	1,61	Early-Late	T12
ct442	7,51	4,21	7,34	5,33	3,87	1,43	4,97	0,17	17,21	3,26	30,92	2,77	Late	T42
ct449	23,71	10,70	19,48	9,99	11,41	2,43	4,82	2,35	6,30	0,78	24,34	2,49	Early-Late	T42
ct483	19,74	2,33	18,46	4,62	20,82	4,87	16,57	0,95	11,23	0,11	21,35	1,19	Early	T42
ct484	9,78	ND	12,21	ND	9,74	1,85	10,79	1,50	9,28	2,63	17,60	0,94	Early	T42
ct565	43,37	5,73	28,20	5,11	27,65	1,30	111,13	43,30	284,17	5,01	398,46	47,97	Late	T42
ct618	456,52	19,28	296,09	19,93	235,44	8,84	86,28	1,71	30,12	1,27	36,26	3,96	Early	T2
ct728	13,24	ND	45,79	9,90	80,28	3,66	71,31	8,02	42,94	3,60	34,59	6,79	Mid	T12
ct789	25,74	3,82	28,47	5,15	24,34	0,91	23,54	1,79	13,01	0,17	24,22	1,24	Early	T6
ct813	5,08	1,30	4,74	1,82	6,50	0,23	3,72	0,11	31,28	15,74	56,08	31,18	Late	T42
ct850	5,40	0,87	8,88	1,29	24,04	4,05	14,64	1,64	11,56	5,00	17,60	3,26	Mid	T12

[a]Real-time quantitative PCR (RT-qPCR) of mRNA levels of *inc* genes during the developmental cycle of *Chlamydia trachomatis* E/Bour strain. The data shows for each *inc* at each time point the mean values ($x10^5$) ± SEM, normalized against the corresponding values of the *16SrRNA*, from at least two independent experiments (except where indicated not determined, ND).

Table A.10 – mRNA levels of *inc* genes during the developmental cycle of *Chlamydia trachomatis* L2/434 strain[a].

***Chlamydia trachomatis* L2/434**														
	Mean and SEM values at different time points (h) post-infection (p.i.)												**Profile of expression**	**Peak of expression (time p.i)**
inc	**T2**		**T6**		**T12**		**T20**		**T30**		**T42**			
	Mean	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**		
ct005	94,67	37,18	86,71	35,17	24,83	8,20	17,62	6,28	213,31	0,31	174,63	53,33	Early-Late	T30
ct006	106,04	29,67	132,78	ND	27,17	ND	10,53	4,51	8,06	0,68	8,37	2,12	Early	T6
ct036	84,10	11,58	52,81	16,11	11,77	3,43	16,60	3,82	10,13	0,77	9,93	2,54	Early	T2
ct058	28,82	15,47	51,40	30,05	22,64	8,73	10,75	6,61	4,45	0,55	6,83	1,14	Early	T6
ct101	6,63	3,34	27,89	10,76	30,64	7,80	34,08	12,26	7,63	1,77	7,31	0,69	Mid	T20
ct115	258,95	22,03	126,79	1,33	96,86	20,81	62,57	23,85	15,31	0,81	26,46	1,52	Early	T2
ct116	415,04	61,54	238,10	41,33	170,65	25,07	92,53	28,56	28,95	1,05	40,75	10,11	Early	T2
ct117	176,10	22,38	137,97	29,90	151,18	42,07	82,87	29,44	17,14	0,71	22,00	2,46	Early	T2
ct118	163,41	59,99	153,76	45,90	215,10	65,12	109,37	45,41	22,45	3,81	21,43	5,92	Early	T12
ct119	7,26	4,55	4,29	0,02	10,71	1,29	33,35	1,91	14,82	0,62	7,25	2,31	Mid	T20
ct134	369,19	168,94	204,78	85,05	92,14	31,13	33,57	21,44	14,59	1,93	19,70	1,54	Early	T2
ct135	255,73	98,72	162,86	76,44	66,40	25,82	26,76	13,14	12,88	0,21	18,24	3,31	Early	T2
ct147	124,57	9,03	129,62	25,79	68,39	3,50	25,95	0,28	21,89	1,64	27,53	1,93	Early	T6
ct164	9,11	2,71	9,61	1,86	9,82	1,32	11,98	1,11	4,47	0,12	3,74	0,37	Early	T20
ct179	54,71	3,10	79,12	13,31	31,26	7,29	35,44	9,17	23,56	1,80	19,85	6,77	Early	T6
ct192	115,14	13,19	53,71	1,29	16,38	1,15	5,46	1,31	3,21	0,76	6,71	3,34	Early	T2
ct195	18,00	6,59	17,26	2,65	15,52	4,88	16,31	2,72	6,16	0,89	5,49	0,78	Early	T2
ct196	130,15	65,05	42,64	12,60	12,11	0,63	5,85	2,33	4,65	0,20	6,99	2,13	Early	T2
ct214	91,69	18,88	42,99	8,85	13,86	0,63	10,26	2,69	93,49	5,36	103,13	27,14	Early-Late	T42
ct222	2,54	0,62	4,37	1,24	23,41	7,36	16,26	6,44	6,39	0,23	7,67	1,35	Mid	T12
ct223	6,53	2,79	6,15	0,28	96,29	14,28	87,33	35,38	27,76	6,66	18,47	2,62	Mid	T12
ct224	6,21	2,79	10,64	2,24	7,97	1,17	7,18	2,57	3,90	0,19	4,74	1,38	Early	T6
ct225	11,33	5,55	13,39	4,00	17,76	4,30	14,60	5,71	6,14	0,51	6,61	2,00	Early	T12
ct226	13,58	4,89	14,18	5,36	19,34	2,71	13,22	1,41	6,04	1,45	6,24	2,27	Early	T12
ct227	33,46	17,48	27,99	1,34	11,79	2,58	7,44	3,01	3,96	0,10	5,99	2,66	Early	T2
ct228	208,44	55,66	105,75	15,64	31,60	1,89	10,87	3,32	6,86	0,44	12,67	3,80	Early	T2
ct229	758,67	178,59	265,91	46,22	38,12	10,04	14,43	5,33	22,81	0,34	45,61	0,68	Early	T2
ct232	326,18	96,78	294,04	3,21	117,49	18,03	58,85	16,80	18,79	4,01	15,38	0,36	Early	T2
ct233	65,27	34,81	63,78	10,94	36,87	14,49	36,69	13,76	16,25	5,09	11,59	1,63	Early	T2

Chlamydia trachomatis L2/434														
	Mean and SEM values at different time points (h) post-infection (p.i.)												**Profile of expression**	**Peak of expression (time p.i)**
inc	**T2**		**T6**		**T12**		**T20**		**T30**		**T42**			
	Mean	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**		
ct249	113,93	26,16	72,74	2,54	30,02	6,54	21,43	6,92	10,38	0,93	14,90	0,54	Early	T2
ct288	195,11	91,23	113,81	20,10	27,98	2,85	15,59	3,09	137,11	34,38	85,22	22,50	Early-Late	T2
ct300	17,29	4,10	31,93	8,34	21,71	4,57	19,12	2,41	16,79	0,78	29,37	2,75	Early	T6
ct345	4,56	1,12	10,86	0,52	13,81	0,40	17,20	1,49	10,20	0,81	5,44	0,56	Mid	T20
ct357	71,54	12,84	45,51	7,05	21,15	3,41	18,54	1,58	43,63	1,15	58,15	12,22	Early-Late	T2
ct358	56,45	23,45	53,88	10,20	35,93	6,57	25,71	5,54	50,13	15,28	42,76	8,79	Early-Late	T2
ct365	11,76	ND	8,06	1,89	5,18	0,90	9,76	1,82	32,44	3,33	26,03	1,96	Late	T30
ct383	95,36	20,18	57,68	6,91	26,40	3,87	20,72	ND	12,00	2,79	8,50	0,98	Early	T2
ct440	47,24	20,92	49,31	3,13	15,33	1,78	9,07	2,25	4,80	0,37	6,29	1,37	Early	T6
ct442	16,75	6,87	11,08	0,30	6,65	1,19	23,95	2,91	40,16	11,28	31,69	1,52	Late	T30
ct449	94,37	43,09	55,11	4,62	17,68	3,96	8,95	1,51	5,23	0,70	8,28	1,44	Early	T2
ct483	23,22	4,77	18,12	3,45	16,42	1,15	19,35	1,96	9,53	0,33	12,09	1,16	Early	T2
ct484	10,50	3,18	8,68	3,47	13,39	1,60	24,26	5,18	8,07	1,63	7,17	0,76	Early	T20
ct565	39,26	0,79	23,72	0,91	22,69	3,98	70,41	18,78	164,61	22,37	128,83	24,22	Late	T30
ct618	282,40	84,17	324,18	21,24	229,60	9,71	130,26	14,22	27,23	0,87	34,17	3,52	Early	T6
ct728	40,24	7,34	79,25	22,97	82,33	4,69	76,08	7,33	30,77	1,47	15,16	1,54	Mid	T12
ct789	57,27	10,47	38,58	8,60	24,81	3,29	23,17	3,37	8,33	2,92	7,45	0,76	Early	T2
ct813	7,53	2,32	9,64	1,20	15,86	4,34	8,87	0,83	51,37	25,19	45,72	13,93	Late	T30
ct850	6,22	3,52	10,17	3,77	19,20	5,72	23,86	5,65	12,52	0,98	8,23	1,53	Mid	T20

[a]Real-time quantitative PCR (RT-qPCR) of mRNA levels of *inc* genes during the developmental cycle of *Chlamydia trachomatis* L2/434 strain. The data shows for each *inc* at each time point the mean values ($x10^5$) ± SEM, normalized against the corresponding values of the *16SrRNA*, from at least two independent experiments (except where indicated not determined, ND).

Table A.11 – mRNA levels of *inc* genes during the developmental cycle of *Chlamydia trachomatis* B/Har36 strain[a].

***Chlamydia trachomatis* B/Har36**														
	Mean and SEM values at different time points (h) post-infection (p.i.)												**Profile of expression**	**Peak of expression (time p.i)**
inc	**T2**		**T6**		**T12**		**T20**		**T30**		**T42**			
	Mean	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**		
ct005	37,75	5,34	31,71	9,02	18,79	1,25	11,02	2,27	5,48	2,11	90,67	18,09	Late	T42
ct006	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct036	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct058	0,03	0,01	0,01	0,01	0,03	0,00	0,08	0,04	0,02	0,01	0,01	0,00	Mid	T20
ct101	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct115	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct116	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct117	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct118	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct119	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct134	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct135	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct147	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct164	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct179	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct192	6,13	2,47	12,42	6,27	10,22	0,92	8,33	2,02	2,27	0,86	1,82	0,51	Mid	T6
ct195	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct196	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct214	30,72	17,70	20,81	8,67	17,00	1,80	13,93	0,13	16,32	3,11	78,44	22,59	Late	T42
ct222	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct223	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct224	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct225	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct226	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct227	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct228	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct229	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct232	49,32	22,66	58,56	28,45	56,30	11,84	59,13	11,51	17,65	8,47	9,28	3,06	Early	T20
ct233	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND

***Chlamydia trachomatis* B/Har36**														
	Mean and SEM values at different time points (h) post-infection (p.i.)												**Profile of expression**	**Peak of expression (time p.i)**
inc	**T2**		**T6**		**T12**		**T20**		**T30**		**T42**			
	Mean	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**		
ct249	25,11	9,55	20,91	6,97	17,70	4,67	26,63	1,84	11,98	2,64	8,48	0,23	Early	T20
ct288	12,25	3,24	10,49	4,50	11,13	3,34	8,79	0,56	5,14	1,61	33,15	4,85	Late	T42
ct300	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct345	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct357	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct358	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct365	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct383	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct440	10,48	3,18	11,37	3,78	13,61	0,77	14,00	3,59	7,06	2,61	3,98	1,21	Early	T20
ct442	4,89	0,02	4,20	1,04	4,79	0,98	22,61	0,76	15,12	5,82	14,61	4,71	Mid	T20
ct449	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct483	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct484	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct565	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct618	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct728	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct789	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct813	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct850	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND

[a]Real-time quantitative PCR (RT-qPCR) of mRNA levels of *inc* genes during the developmental cycle of *Chlamydia trachomatis* B/Har36 strain. The data shows for each *inc* at each time point the mean values ($x10^5$) ± SEM, normalized against the corresponding values of the *16SrRNA*, from at least two independent experiments (except where indicated not determined, ND).

Table A.12 – mRNA levels of *inc* genes during the developmental cycle of *Chlamydia trachomatis* F/CS465-95 strain[a].

***Chlamydia trachomatis* F/CS465-95**														
	Mean and SEM values at different time points (h) post-infection (p.i.)												**Profile of expression**	**Peak of expression (time p.i)**
inc	**T2**		**T6**		**T12**		**T20**		**T30**		**T42**			
	Mean	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**		
ct005	55,95	5,01	31,64	2,27	25,35	2,17	12,96	0,46	160,43	9,03	190,35	1,84	Late	T42
ct006	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct036	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct058	0,06	0,03	0,01	0,01	0,05	0,05	0,03	0,00	0,02	0,00	0,01	0,01	Early	T2
ct101	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct115	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct116	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct117	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct118	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct119	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct134	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct135	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct147	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct164	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct179	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct192	3,32	0,39	2,47	1,05	2,59	0,77	1,86	0,03	1,98	0,57	1,70	0,95	Early	T2
ct195	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct196	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct214	14,04	0,32	10,36	0,10	6,27	0,76	5,64	0,55	53,56	1,50	36,15	4,04	Late	T30
ct222	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct223	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct224	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct225	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct226	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct227	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct228	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct229	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct232	34,84	11,68	25,41	2,24	27,00	3,38	25,28	1,10	10,86	2,19	6,21	2,06	Early	T2
ct233	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND

***Chlamydia trachomatis* F/CS465-95**														
	Mean and SEM values at different time points (h) post-infection (p.i.)												**Profile of expression**	**Peak of expression (time p.i)**
inc	**T2**		**T6**		**T12**		**T20**		**T30**		**T42**			
	Mean	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**		
ct249	27,93	8,62	22,31	1,78	16,35	1,75	12,24	1,75	8,29	0,55	9,01	2,25	Early	T2
ct288	57,61	12,95	29,04	3,24	32,01	8,06	12,82	3,36	122,50	3,27	115,49	30,88	Late	T30
ct300	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct345	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct357	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct358	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct365	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct383	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct440	6,85	0,72	6,99	0,60	8,81	0,88	7,83	0,34	5,25	0,94	3,05	0,85	Early	T12
ct442	3,61	1,08	3,16	1,26	3,03	1,41	3,58	0,66	12,82	3,07	9,48	3,98	Late	T30
ct449	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct483	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct484	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct565	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct618	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct728	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct789	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct813	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct850	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND

[a]Real-time quantitative PCR (RT-qPCR) of mRNA levels of *inc* genes during the developmental cycle of *Chlamydia trachomatis* F/CS465-95 strain. The data shows for each *inc* at each time point the mean values ($x10^5$) ± SEM, normalized against the corresponding values of the *16SrRNA*, from at least two independent experiments (except where indicated not determined, ND).

Table A.13 – mRNA levels of *inc* genes during the developmental cycle of *Chlamydia trachomatis* L2b/CS19-08 strain[a].

***Chlamydia trachomatis* L2b/CS19-08**														
	Mean and SEM values at different time points (h) post-infection (p.i.)												**Profile of expression**	**Peak of expression (time p.i)**
Inc	**T2**		**T6**		**T12**		**T20**		**T30**		**T42**			
	Mean	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**		
ct005	50,19	5,53	45,67	6,89	22,54	7,02	14,11	5,24	171,04	59,72	260,88	14,60	Late	T42
ct006	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct036	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct058	8,11	0,67	9,78	2,93	10,02	4,34	4,55	1,02	2,19	0,77	6,04	3,03	Early	T12
ct101	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct115	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct116	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct117	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct118	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct119	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct134	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct135	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct147	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct164	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct179	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct192	39,84	3,67	30,38	7,39	12,12	0,42	6,69	1,26	3,71	0,11	10,61	2,95	Early	T2
ct195	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct196	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct214	79,07	13,49	83,27	48,40	23,17	5,16	15,75	0,37	117,14	2,71	146,67	48,17	Early-Late	T42
ct222	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct223	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct224	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct225	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct226	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct227	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct228	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct229	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct232	118,54	32,04	181,11	58,21	184,56	75,39	79,82	17,97	25,24	7,86	40,13	18,54	Early	T12
ct233	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND

***Chlamydia trachomatis* L2b/CS19-08**														
	Mean and SEM values at different time points (h) post-infection (p.i.)												**Profile of expression**	**Peak of expression (time p.i)**
Inc	**T2**		**T6**		**T12**		**T20**		**T30**		**T42**			
	Mean	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**		
ct249	38,21	13,32	31,45	9,09	16,70	5,97	11,87	1,75	7,88	3,55	14,94	7,11	Early	T2
ct288	14,14	3,09	10,59	2,39	7,09	2,25	10,12	2,89	61,11	30,63	56,41	14,82	Late	T30
ct300	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct345	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct357	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct358	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct365	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct383	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct440	13,18	3,01	18,71	1,77	9,03	1,38	6,87	1,06	4,28	0,11	10,74	2,19	Early	T6
ct442	7,45	1,00	8,15	5,53	5,18	3,35	18,28	4,37	29,10	10,93	27,00	7,43	Late	T30
ct449	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct483	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct484	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct565	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct618	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct728	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct789	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct813	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct850	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND

[a]Real-time quantitative PCR (RT-qPCR) of mRNA levels of *inc* genes during the developmental cycle of *Chlamydia trachomatis* L2b/CS19-08 strain. The data shows for each *inc* at each time point the mean values ($x10^5$) ± SEM, normalized against the corresponding values of the *16SrRNA*, from at least two independent experiments (except where indicated not determined, ND).

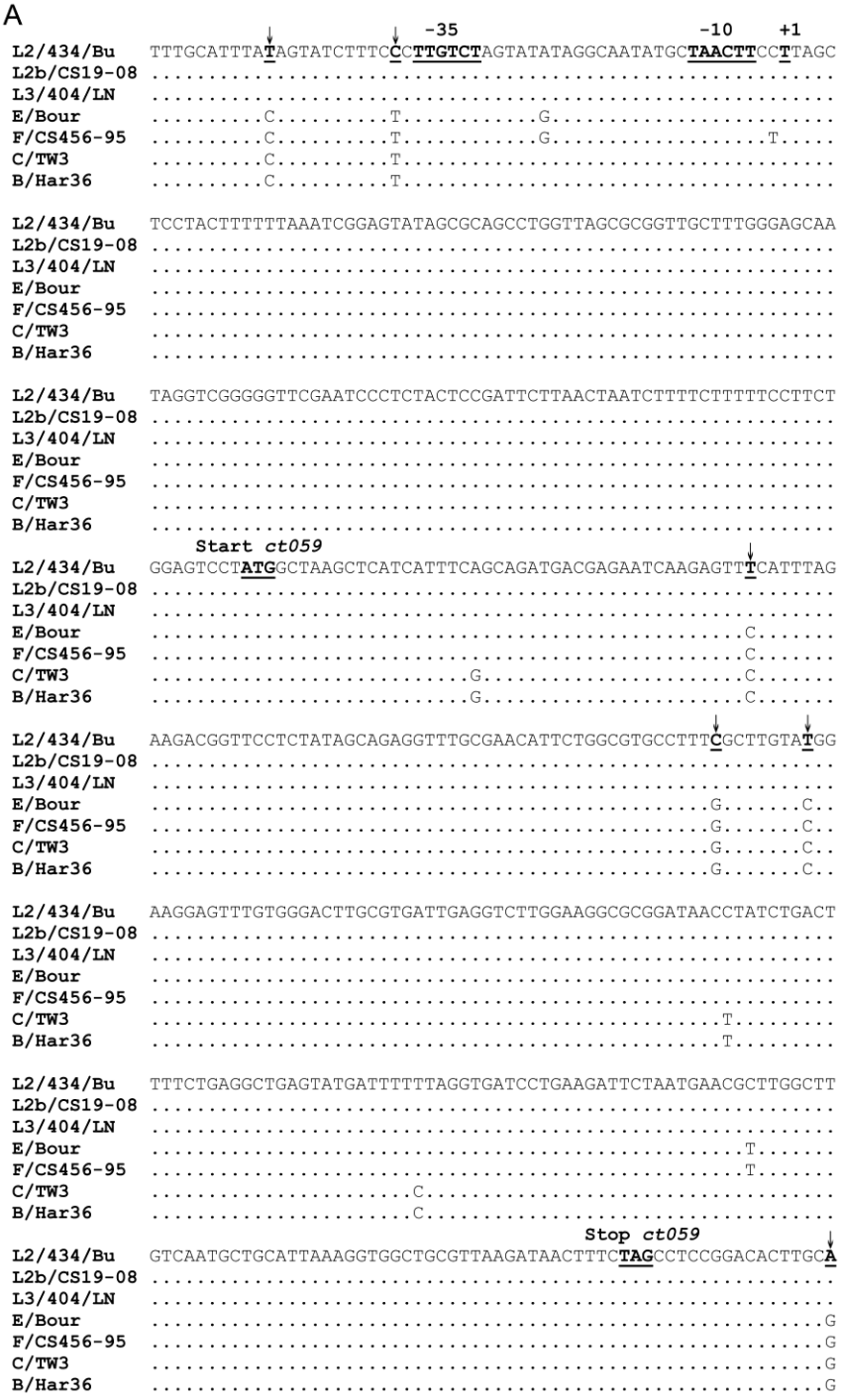
Table A.14 – mRNA levels of *inc* genes during the developmental cycle of *Chlamydia trachomatis* L3/404 strain[a].

***Chlamydia trachomatis* L3/404**														
	Mean and SEM values at different time points (h) post-infection (p.i.)												**Profile of expression**	**Peak of expression (time p.i)**
Inc	**T2**		**T6**		**T12**		**T20**		**T30**		**T42**			
	Mean	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**		
ct005	36,91	0,77	31,50	0,35	14,88	2,06	27,19	2,20	144,46	9,06	103,73	7,90	Late	T30
ct006	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct036	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct058	4,76	2,33	9,49	1,22	11,03	3,13	3,82	0,20	2,69	1,15	3,22	1,68	Early	T12
ct101	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct115	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct116	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct117	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct118	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct119	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct134	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct135	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct147	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct164	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct179	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct192	30,78	4,79	23,57	4,00	10,75	0,53	4,60	0,18	3,60	1,65	4,17	1,32	Early	T2
ct195	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct196	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct214	69,48	3,03	63,39	11,58	31,98	2,07	43,91	2,80	156,14	56,50	83,04	10,85	Early-Late	T30
ct222	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct223	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct224	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct225	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct226	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct227	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct228	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct229	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct232	109,37	25,18	168,47	45,40	135,46	16,09	62,16	0,51	18,57	0,01	24,92	4,48	Early	T6
ct233	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND

***Chlamydia trachomatis* L3/404**														
	Mean and SEM values at different time points (h) post-infection (p.i.)												**Profile of expression**	**Peak of expression (time p.i)**
Inc	**T2**		**T6**		**T12**		**T20**		**T30**		**T42**			
	Mean	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**	**Mean**	**SEM**		
ct249	68,13	22,47	45,22	2,76	25,77	4,18	15,92	0,44	10,67	2,43	12,38	2,18	Early	T2
ct288	70,34	2,46	48,77	12,40	20,54	2,40	43,77	5,00	101,45	7,00	72,16	1,15	Early-Late	T30
ct300	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct345	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct357	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct358	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct365	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct383	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct440	16,86	5,83	16,50	4,36	12,27	0,96	7,81	0,95	4,41	1,35	4,73	1,76	Early	T2
ct442	15,70	1,79	15,13	4,98	7,43	0,19	23,01	1,18	29,95	26,78	33,04	3,63	Late	T42
ct449	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct483	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct484	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct565	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct618	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct728	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct789	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct813	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
ct850	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND

[a]Real-time quantitative PCR (RT-qPCR) of mRNA levels of *inc* genes during the developmental cycle of *Chlamydia trachomatis* L3/404 strain. The data shows for each *inc* at each time point the mean values ($x10^5$) ± SEM, normalized against the corresponding values of the *16SrRNA*, from at least two independent experiments (except where indicated not determined, ND).

A


-35
-10
+1
L2/434/Bu TTTGCATTTATAGTATCTTTCCCTTGTCTAGTATATAGGCAATATGCTAACTTCCTTAGC
L2b/CS19-08
L3/404/LN
E/Bour .....C.....T.....G.....
F/CS456-95 .....C.....T.....G.....T.....
C/TW3 .....C.....T.....
B/Har36 .....C.....T.....
L2/434/Bu TCCTACTTTTTTAAATCGGAGTATAGCGCAGCCTGGTTAGCGCGGTTGCTTTGGGAGCAA
L2b/CS19-08
L3/404/LN
E/Bour
F/CS456-95
C/TW3
B/Har36
L2/434/Bu TAGGTCGGGGGTTCGAATCCCTCTACTCCGATTCTTAACTAATCTTTTCTTTTTCCTTCT
L2b/CS19-08
L3/404/LN
E/Bour
F/CS456-95
C/TW3
B/Har36
Start ct059
L2/434/Bu GGAGTCCTATGGCTAAGCTCATCATTTCAGCAGATGACGAGAATCAAGAGTTTCATTTAG
L2b/CS19-08
L3/404/LN
E/Bour .....C.....
F/CS456-95 .....C.....
C/TW3 .....G.....C.....
B/Har36 .....G.....C.....
L2/434/Bu AAGACGGTTCCTCTATAGCAGAGGTTTGCGAACATTCTGGCGTGCCTTTCGCTTGTATGG
L2b/CS19-08
L3/404/LN
E/Bour .....G.....C..
F/CS456-95 .....G.....C..
C/TW3 .....G.....C..
B/Har36 .....G.....C..
L2/434/Bu AAGGAGTTTGTGGGACTTGCGTGATTGAGGTCTTGGAAGGCGCGGATAACCTATCTGACT
L2b/CS19-08
L3/404/LN
E/Bour
F/CS456-95
C/TW3 .....T.....
B/Har36 .....T.....
L2/434/Bu TTTCTGAGGCTGAGTATGATTTTTTAGGTGATCCTGAAGATTCTAATGAACGCTTGGCTT
L2b/CS19-08
L3/404/LN
E/Bour .....T.....
F/CS456-95 .....T.....
C/TW3 .....C.....
B/Har36 .....C.....
Stop ct059
L2/434/Bu GTCAATGCTGCATTAAAGGTGGCTGCGTTAAGATAACTTTCTAGCCTCCGGACACTTGCA
L2b/CS19-08
L3/404/LN
E/Bour .....G
F/CS456-95 .....G
C/TW3 .....G
B/Har36 .....G

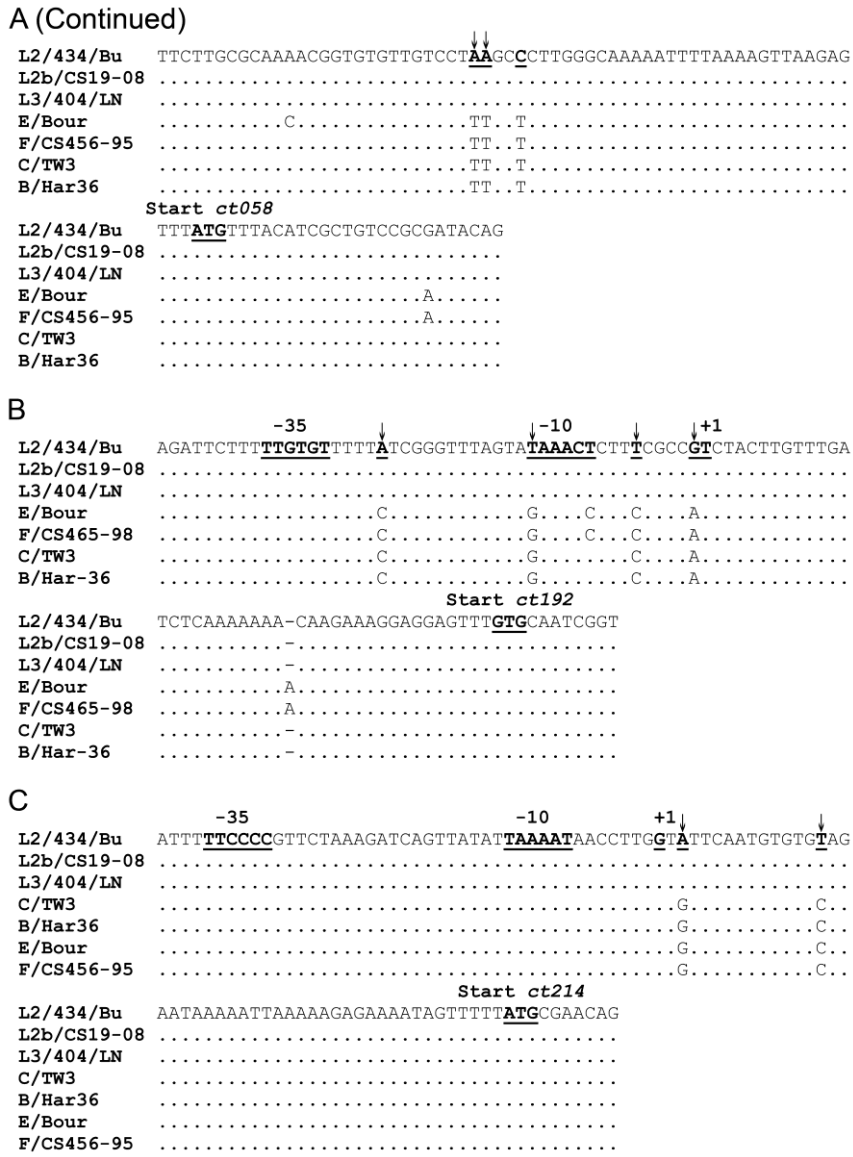


Figure A.1 – LGV-specific nucleotide differences in the promoter region of *ct059*, coding sequence of *ct059*, *ct059-ct058* intragenic region, and first codons of *ct058* (A), the promoter region of *ct192* (B) and of *ct214* (C). The nomenclature of D/UW3 is used. A dot "." indicates an identical nucleotide among all strains. The following features of strain L2/434 are depicted in the sequence (underlined and bold): the transcriptional start site of *ct059*, *ct192*, and *ct214* determined by RACE (+1), the predicted start and stop codons of *ct059*, the deduced -10 and -35 hexamers within the promoter of *ct059*, *ct192*, and *ct214*, the predicted start codon of *ct058*, *ct192*, and *ct214*, and LGV specific nucleotide differences (arrows).